To reduce file size and facilitate electronic distribution, this document has been separated into two files. This is the first of two files; please refer to Appendix 3C – Part II for the remainder of the document.

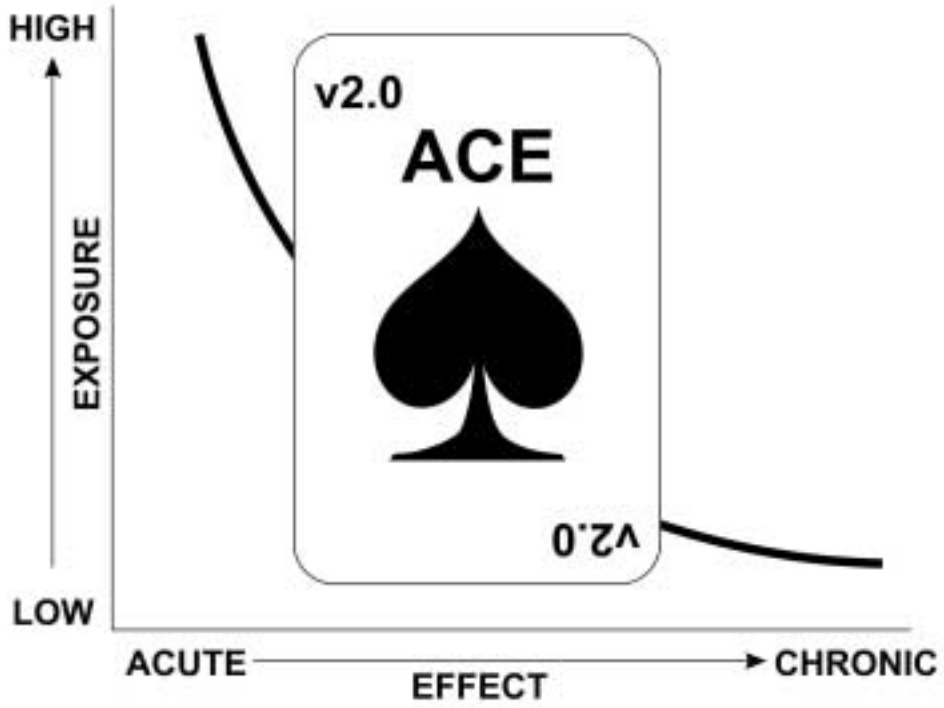
Appendix 3C
Software documentation

Acute-to-Chronic Estimation (ACE v 2.0)
BurrliOZ Help, Licensing, Readme file
Level I Fugacity Model License, Description, Update
Level II Fugacity Model Licence, Description, Update
Level III Fugacity Model Licence, Description, Update
ETX 1.3
ETX 2.0
EXAMS
Interspecies Correlation Estimations (ICE v 1.0)

United States Environmental Protection Agency
Office of Research and Development Washington, DC 20460
EPA/600/R-03/107 December 2003

Acute-to-Chronic Estimation (Ace v 2.0) with Time - Concentration - Effect Models

User Manual and Software



EPA/600/R-03/107
December 2003

Acute-to-Chronic Estimation (ACE v 2.0) with Time-Concentration-Effect Models

User Manual and Software

By

Mark R. Ellersieck, Amha Asfaw, Foster L. Mayer*, Gary F. Krause, Kai Sun, and Gunhee Lee

University of Missouri-Columbia
College of Agriculture, Food and Natural Resources
Agricultural Experiment Station-Statistics
Columbia, MO 65211

*U.S. Environmental Protection Agency
Office of Research and Development
National Health and Environmental Effects Research Laboratory
Gulf Ecology Division
Gulf Breeze, FL 32561-5299

U.S. Environmental Protection Agency
Office of Research Development
1200 Pennsylvania Avenue, NW
Washington, DC 20460

Notice

The U.S. Environmental Protection Agency through its Offices of Research and Development, Pesticide Programs, Pollution Prevention and Toxics, and Water partially funded and collaborated in the research described herein under EPA Project No. CR82827901 to University of Missouri-Columbia, College of Agriculture, Food and Natural Resources, Agricultural Experiment Station-Statistics. It has been subjected to the Agency's peer and administrative review and has been approved for publication as an EPA document. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Contents

List of Figures

Abstract

Predictive toxicological models, including estimates of uncertainty, are necessary to address probability-based ecological risk assessments. Methods and software (ACE) were developed for estimating chronic toxicity from raw acute toxicity data (all response observations at all times and exposures). Three methods were developed - - Accelerated Life Testing (ALT), Multifactor Probit Analysis (MPA), and two-stage Linear Regression Analysis (LRA). Of the three, the method of choice is ALT, in that time to failure (death) of each experimental unit is independent. It requires three partial responses over the time period of acute testing, but will function with one. The MPA is a two dimensional probit analysis using both time and concentration to produce a multiple regression equation, however, each experimental unit is not independent. Also, the MPA requires more partial responses than the ALT. The LRA calculates LC values for each time period and then regresses the LC values as the Y axis and the reciprocal of time as the X axis. The Y intercept is the chronic no-effect concentration. The LRA will function when ALT and MPA fail; no partial responses are required. All methods provide confidence limits for the point estimates. The methods have previously been shown to estimate chronic no-effect concentrations very well when validated against actual paired acute and chronic test results with fishes.

Introduction

Both understanding and evaluating chronic toxicity of chemicals are essential to assessing their ecological hazards and making environmentally sound management decisions. Because of the large number and variety of industrial, agricultural and home-use chemicals released in the U.S. annually and the high cost and effort required for chronic tests, resources are often insufficient to obtain experimental information about long-term environmental impacts for all potentially hazardous chemicals. In comparison, acute tests are less costly and time consuming and, for these reasons, an abundance of acute toxicity data exists for numerous chemicals and organisms. Also, procedures have been developed for extrapolating effects data within classes of chemicals sharing similar chemical structures (Lipnick 1995). Thus, there is a strong rationale to relate acute and chronic toxicities of chemicals and to develop statistical and mathematical techniques to predict chronic toxicity based on data from acute experiments.

Use of short-term tests as a basis for linkage of exposure and time to response with chronic effects for ecological risk assessments is significant. The ability to accurately and precisely associate chronic effects from acute time-concentration-effect data is a powerful approach that integrates various aspects of toxicokinetics and directly addresses a variety of uncertainties in terms of chronicity. Three models were developed (Lee et al. 1995; Mayer et al. 1994, 2002; Sun et al. 1995b), tying together classical methods (e.g., probit regression) (Finney 1978) and time to event methods (Newman 1994) to provide models that predict chronic toxicity from acute toxicity data.

- Accelerated Life Testing (ALT) – A survival analysis and population-based approach (Weibull distribution) using accelerated life testing theory (Mayer et al. 2002, Sun et al. 1995b). The method was originally used for mechanical and electrical devices placed under short-term or "acute" stress (e.g., generator running constantly at full power and high heat) to predict long-term or "chronic" time to failure. In the ACE software, the model is applied to organisms placed under acute stress (i.e., toxicant), and the variable measured is time to failure or death. The model assumes that both exposure concentrations and duration affect survival probability, and hence, has the ability to summarize the entire concentration-time-response data of a toxicity test. Actual proportion responses are used; probit transformations are not applied. ALT also takes into account the spontaneous survival probability and is suitable to describe both acute and chronic lethality data. The survival function includes competing risks, with contaminant exposure being one.

- Multifactor Probit Analysis (MPA) – Multiple regression models that simultaneously evaluate the relationship among exposure concentration, time, and probit % mortality to predict chronic response (Mayer et al. 2002, Lee et al. 1995). This model is appropriate when different experimental units are present for concentration-time combinations (i.e., one complete replicate is removed at one or more time intervals for a measurement different than survival; only the remaining replicates are used for the remainder of the toxicity test). ALT and LRA models are more appropriate for predicting chronicity from standard acute toxicity data; however, multiple regression models, such as MPA, are necessary when estimating chronicity under changing conditions (e.g., varying exposure scenarios in effluents).

- Linear Regression Analysis (LRA) – A two-step linear regression analysis (Mayer et al. 1994, Mayer et al. 2002). This model combines two linear regressions: 1) estimates low lethal concentrations at each observation time period and 2) regresses those concentrations (dependent variable) against the reciprocal of time (independent variable), with the intercept being the chronic no-effect concentration. Probit transformations of percent response are used.

The software program, Acute-to-Chronic Estimation (ACE), described herein, allows the user to estimate chronic toxicity for a species from raw acute toxicity data with accuracy and precision. ACE will, therefore, greatly enhance the use of probability-based risk assessments for chemicals having minimal data sets. However, if a chronic test is to be conducted, ACE can be used to more accurately identify the range of exposure concentrations required. ACE is based on the Windows platform and is specifically designed

for estimating chronic toxicity and providing graphical and tabular presentation of results. ACE v 2.0 is an upgrade of the former DOS version (Mayer et al. 1999).

Background

Using acute mortality data to estimate chronic toxicity (survival, growth, reproduction) to aquatic organisms customarily involves deriving an application factor (Mount and Stephan 1967) or an acute-to-chronic ratio (Kenaga 1982), both of which require acute and chronic toxicity testing. Kenaga (1979) reviewed the principal measurements of the acute LC50, the maximum acceptable toxicant concentration (MATC), and the application factor (AF) used in determining chronic NOECs (highest concentration causing 0% or no statistically significant effect) for many chemicals. The AF is derived by dividing the MATC for a compound, as determined in a chronic toxicity test with a given species, by the acute LC50 for the same compound tested with the same species. The acute-to-chronic ratio (ACR) is the inverse of the AF. The AF or ACR is then used to estimate chronic NOECs for other species for which only acute toxicity data (EC or LC50s) exist (Buikema et al. 1982). These approaches have limitations.

One limitation is that the biological endpoints and degrees of responses are often not comparable between acute and chronic toxicity data. When either the AF or ACR is used, the acute median lethal concentration (EC or LC50) is compared with the MATC, often derived from an endpoint other than mortality. Although different degrees of response (acute 50% vs. chronic no-effect) could be used when response slopes are similar, the slopes may be different. Additionally, use of the AF or ACR method does not take into consideration the progression of mortality through time that is derived in acute toxicity tests. The concentration-time-response interaction has been addressed by Shirazi and Lowrie (1988), but they directed their efforts toward better defining the LC50. The acute toxicity value represents only one point in time (e.g., 96-h LC50), and the relationship of degree of response with duration of exposure should be essential when chronic toxicity is predicted from acute toxicity data.

Lethality and other toxic effects are dependent on both concentration of a chemical to which an organism is exposed and length of exposure time. It is a common practice to investigate the toxicity of new and existing chemicals and effluents using acute toxicity tests. This is done by observing mortality resulting from exposure to a series of chemical concentrations, usually at 24, 48, 72, and 96 h. Time course distinguishes acute from chronic toxicity and also relates them as an integrated and progressive process. A time to response approach gives a better understanding of the progression of toxic effects over time, and survival time modeling has shown great applicability in toxicological studies (Crane et al. 2002, Dixon and Newman 1991, Newman and Aplin 1992).

The models included here are more comprehensive approaches to predicting chronicity, both toxicologically and statistically. Simultaneous consideration is given to exposure concentration, degree of response, and time course of effect, all of which are usually included in describing the results of an acute toxicity test, but are seldom used in hazard assessment. A consistent endpoint (mortality) and degree of response (~0%) are used to predict long-term (chronic) lethality from acute toxicity test data. These calculations are based solely on raw acute toxicity test data and do not require conducting a chronic toxicity test. Estimated long-term (chronic) lethality values have previously been validated for accuracy with actual chronic no-effect values derived for 28 chemical-fish species combinations (Mayer et al. 2002).

Software Language

The ACE software is based on a Windows® platform and written in Visual Basic (Microsoft® Visual Basic 6.0 1987-2000). Subroutines (Fortran programs) in Visual Basic and Visual Fortran are required to call Fortran IMSL Routines necessary in certain calculations (Compaq Fortran 1999, Visual Numeric 1999).

Installing ACE

System Requirements

- Operates on Microsoft Windows 95, 98, 2000, NT and XP (Windows® 98 or later is suggested).
- Minimum 16 MB RAM (64 MB or greater is suggested).
- CPU speed of over 200 MHz is suggested; ACE will work with less, but is very slow.
- 6MB hard disk space.
- Mouse or pointing device.
- Printer (optional).

Remove any existing versions of ACE before installing the new one or malfunctions may occur.

To remove old ACE software:

1. Double click **My Computer**.
2. Double click **Control Panel**.
3. Double click **Add/Remove Programs**.
4. Click **ACE**.
5. Click **Delete** or **Change/Remove**.

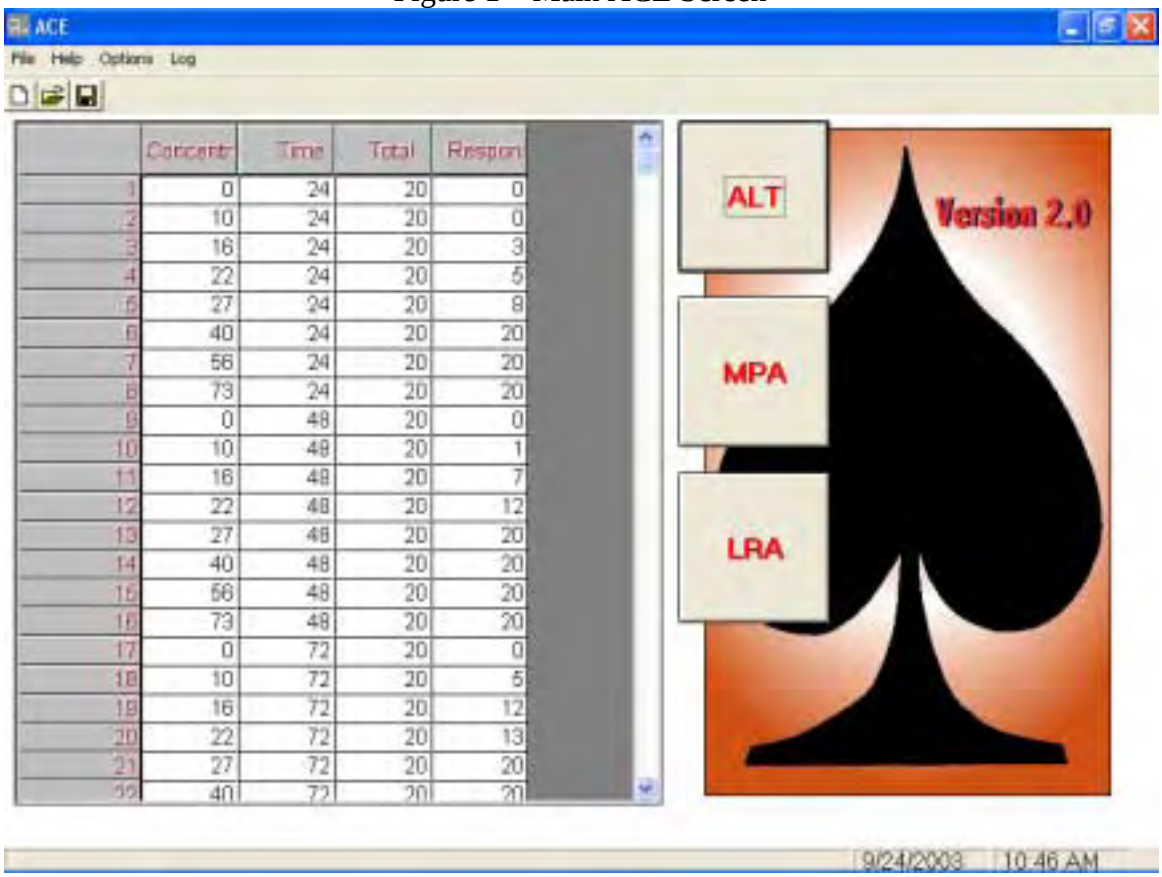
To install new ACE software:

1. Place the ACE CD in the CD ROM drive.
2. Click **Start** button.
3. Select **Run** from the menu.
4. Select **Browse** from the **Run** window.
5. Select drive letter associated with the CD drive from **Browse** window (or ACE 2003 [D:]).
6. Double-click **Setup** file or **D:\SETUP.EXE** file.
7. Click **OK**.
8. Windows now walks you through the installation process. If a "Yes" or "No" question is encountered, choose "Yes".
9. Following installation, the ACE program can be accessed by clicking **Start**, **Programs**, and then **ACE**. You can create an icon on the Desktop screen by placing the mouse pointer on the ACE icon, holding down on the control button, and dragging the icon to desired location on the screen.

Using ACE in Windows

Double click on the ACE icon in the Desktop screen and the main ACE screen will appear (Fig. 1). There are three main sections to the screen. The first section (left) is for data entry or for including data from other sources (e.g., Excel, Lotus 123, etc.). The second section (right center) represents the models available in ACE (ALT, accelerated life testing; MPA, multifactor probit analysis; LRA, linear regression analysis). The third section is the ACE logo, appearing in the background at right. Following data entry and conversion to ASCII files (see below), click on the box for the model of choice (ALT, MPA, LRA), and the analysis results and graphics will automatically be generated.

Figure 1 – Main ACE Screen



Menu Bar - Main Screen

File – Clicking on **File** provides the following drop down menu:

- **New** – Clears spreadsheet so new data can be entered.
- **Open** – Obtains a saved data set from an outside source (see **Obtaining Data from an Outside Source**).
- **Save** – Saves any changes back to the same file name.
- **Save As** – Saves a data set for the first time or saves an existing data set to a new file name.
- **Exit** – Clicking on **Exit** will end the ACE program; clicking on **X** in the upper right-hand corner of the main ACE window will perform the same function as **Exit**.
- Help – User manual.

Options – Option screen will appear; see **OPTIONS** for explanation.
Log – If the ACE program does not run, then an error list will appear; the screen will be empty if no problems occur.
Sheet icon – This is the same as **New** under the **File** drop down menu.
File icon – This is the same as **Open** under the **File** drop down menu.
Floppy disk icon – This is the same as **Save** under the **File** drop down menu.

Menu Bar - ALT, MPA, LRA

- **Print** – Allows printing of selected output (statistical output, graph, or log).
- **Save_on_file** – Saves the statistical output to a file; this is the same as **Save as** described previously.

- **Log** – Provides additional statistical output information.

Data Entry

Format

The following acute toxicity data set for Kepone (Buckler et al. 1981) is used to demonstrate data formatting. The data must be entered in column format as follows, except that columns may be in any order; each column is identified by column headers in the first window (Fig. 1). Data must be entered in the following format for rows:

Concentration	Time (h)	Total (# of Organisms Tested)	Response (# Dead)
0	24	20	0
10	24	20	0
16	24	20	3
22	24	20	5
27	24	20	8
40	24	20	20
56	24	20	20
73	24	20	20
0	48	20	0
10	48	20	1
16	48	20	7
22	48	20	12
27	48	20	20
40	48	20	20
56	48	20	20
73	48	20	20
0	72	20	0
10	72	20	5
16	72	20	12
22	72	20	13
27	72	20	20
40	72	20	20
56	72	20	20
73	72	20	20
0	96	20	0
10	96	20	5
16	96	20	12
22	96	20	13
27	96	20	20
40	96	20	20
56	96	20	20
73	96	20	20

Entering Data Directly

Acute toxicity data can be entered directly to ACE using the spreadsheet (Fig. 1) and keypad functions. The following keypad functions are operational in the spreadsheet: arrow keys, Delete key, Enter key

(functions the same as the down arrow key), and number keys. Each column has to be identified for the ACE program to function properly. Click on each of the column headers, click on arrow, and select appropriate descriptor for that column.

- ID – This is not necessary if a single data set is entered. If more than one data set is to be entered, see **Entering Data from Outside Source** below.
- Concentration – Exposure concentration or % effluent (for extremely large numbers, convert to next higher unit [e.g., µg to mg]).
- Time – Observation time in hours, usually 24, 48, 72, and 96 hours (maximum times are 12).
- Total – Number of organisms exposed per concentration.
- Response – Number of organisms dead or affected.

The ACE default order of column designation is the same as above.

Next, enter the data, click on **File** and then **Save as** and enter a data set name in the file name box. The data set will be saved as a tab delimited file unless an extension name of CSV is typed. An extension name of CSV will results in a comma delimited file. The Tab or Comma delimited file types are preferred. The data are brought back into ACE by clicking on the icon file, data set to be analyzed, and **Open**. Then click on the model of preference (ALT, MPA, LRA), and the analysis is automatically conducted. If data are not analyzed, recheck the column headers to make sure they are correct.

Entering Data from Outside Source

The software is not meant to be a sophisticated spreadsheet, and the best way to enter multiple data sets is from an outside source using softwares capable of producing ASCII text files (e.g., Excel, Word, etc.). If data sets are stacked, a fifth column (ID) must be added in order to identify the different acute data sets.

Once data have been entered, save them as an ASCII file. This is done by clicking on **File** in the upper left corner and then clicking on **Save as**. The **Save as** screen will appear with two boxes at the bottom; **File name** and **Save as type:**. Type in a name for the data set in the File name box. Click **Save as type:**, a list of file types will appear. The following file types are appropriate for the ACE software: **Space delimited, Tab delimited** and **Comma delimited (CSV).** The Tab delimited or CSV file types are preferred.

Obtaining Data from Outside Source

To obtain a data set from an outside source while in the ACE program, click on the File icon in the upper left-hand corner and the following drop down menu will appear:

New Ctrl N
Open
Save
Save As
Exit

Click **Open**; if the data set is not listed in the **Open** screen, click **Files of type:**. Click arrow and then **All(*.*)**. If the file is still not present, click on **Look in:**. This will list all of the disk drives in your computer. Once the data set has been found, double click on the data set and the data will be entered into the ACE program. Again, data sets must be converted to Tab, Comma or Space delimited file types, with Tab and CSV being preferred.

Once the data set is imported into the ACE program in the correct format, title or other descriptive lines must be removed. Click on the line number in the spreadsheet for the line that is to be deleted (left side of main ACE window) and press the Delete key on keyboard.

Each column needs to be identified by the ACE program. Check the column headers on the Main ACE Screen. If they are correct, the program is ready to run. If not, click on each of the column headers and correct (see **Entering Data Directly**).

Data Correction

If data need to be corrected, it can be done within ACE. Just click on the cell, delete the incorrect number with the Delete key, and then correct the entry. If columns are too narrow to fully observe identifiers or numbers, widen the columns by placing the cursor on the right border of the column header and, while holding down the left mouse button, drag to the right until the desired width is achieved. Reverse this process to narrow the columns. Changes are saved by clicking on **File**, selecting either **Save** or **Save as**, entering a name for the data set in **File name:**, and clicking on **Save**.

Model Selection

Brief guidelines for using ACE and selecting the appropriate models are:

1. Exposure Type - Historically, three test exposure techniques have been used to determine acute toxicity for aquatic organisms (static, static renewal, and flow-through). Acute toxicity data used in ACE should be based on static renewal or flow-through techniques, since static exposure may give erroneous results, except for chemicals that are water soluble (see fluridone, Mayer et al. 1994). Further research is needed to determine at what octanol/water or solubility values static test data begin resulting in erroneous chronic predictions.

2. Model Preference – ALT is the method of choice, followed by LRA and MPA, based on experimental designs commonly used in acute toxicity testing. MPA is a special case application and is seldom used.

3. Partial Responses – Dependability of chronicity estimates is generally enhanced with increasing numbers of partial responses (% mortality >0<100%). Recommended partial responses are: ALT $\geq$ 3, MPA $\geq$ 5, and LRA $\geq$ 1. However, ALT will generally function with one partial response; LRA will function with no partial responses as long as there is an exposure-response in time. It is not uncommon to conduct acceptable acute toxicity tests where no partial responses occur, only 0 and 100%; under these conditions, the LRA is the model of choice.

4. Percent Effect for Chronicity – Recommended percent values to be selected for estimated chronic toxicity are: ALT = 1.0%, MPA = 0.01%, and LRA = 0.01%. Use of 0.01% for the MPA and LRA represents a very close approximation to zero on the probit scale (Mayer et al. 1994, Mayer et al. 2002). ALT differs in that 1.0% is presently considered the smallest detectable difference due to the model being population-based (small numbers of organisms usually exposed in each concentration). These percentages correspond well to statistically-based chronic no-effect concentrations for mortality using hypothesis testing (i.e., analysis of variance; Mayer et al. 2002).

ACE Application Windows

Data Analysis

Download a data set to the main ACE screen and click on a model (ALT, MPA, or LRA); the data will automatically be analyzed. Click on the **X** in the upper right hand corner to return to the main screen; a different model can then be selected. When you click on a model on the main screen, a split screen will appear; statistical output on the left and graphics on the right. Double click on either to fill screen; double click again to return to split screen. Click on the **X** in the upper right-hand corner of the main screen to exit ACE.

Printing Output

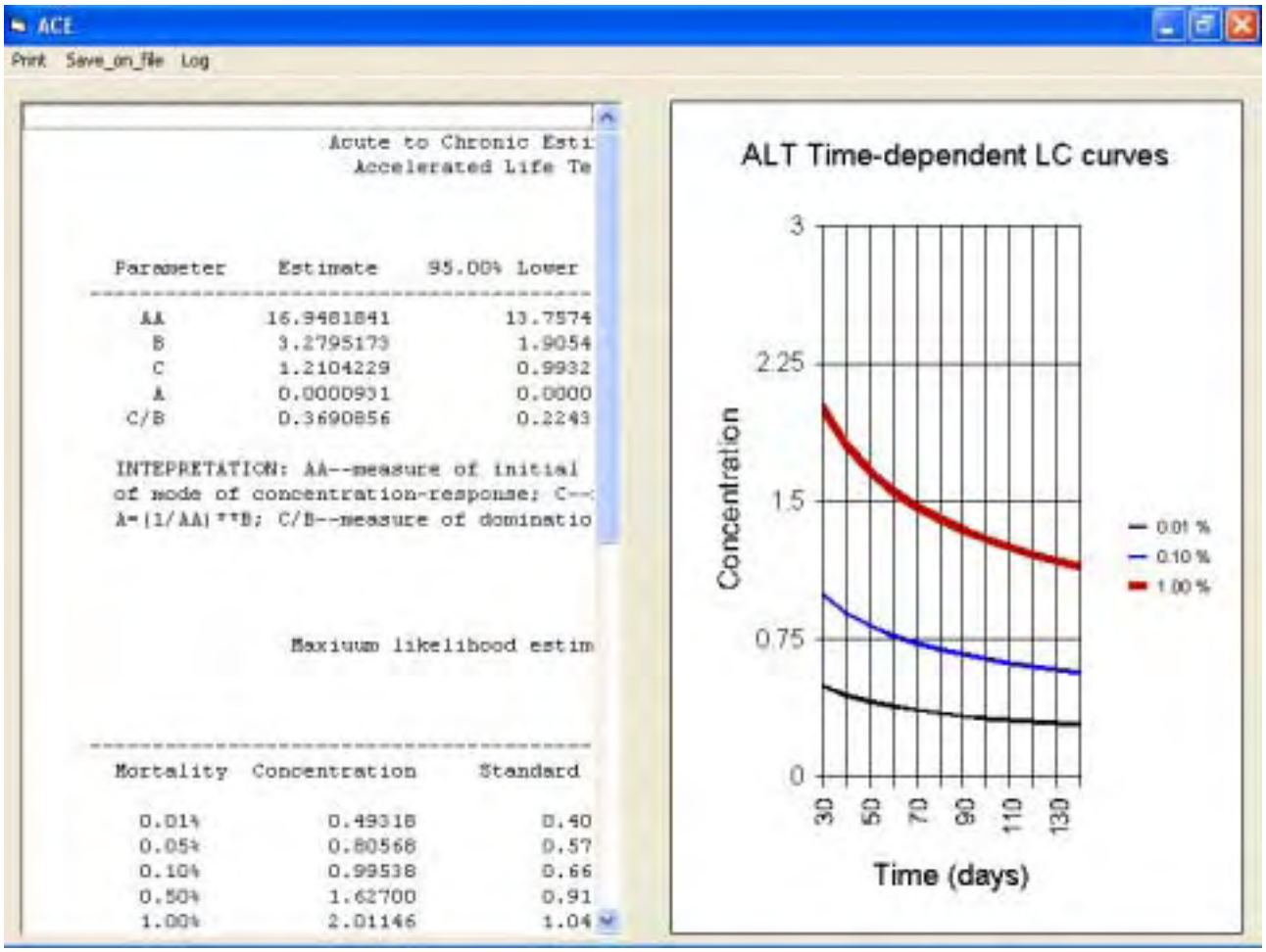
Printing of the statistical or graphics output is achieved by clicking **Print**, or the outputs can be saved by clicking **Save_on_file** (upper left-hand corner of screen). Additional statistical output can be obtained by clicking on **Log**. The output for **Log** includes the statistical output plus the additional information below and can also be printed or saved.

- ALT – Data input, iterations required to solve function estimates, variance-covariance matrix for function estimates to estimate confidence intervals, and data used in the analysis (the highest concentration having 0% response and the lowest concentration having 100% response are used for each observation time).
- MPA – Data used in the analysis as described in ALT.
- LRA – Statistical analyses for all six models including slope, estimated no-effect chronic concentration, confidence intervals, r^2, and data used in the analysis as described in ALT.

ALT- Accelerated Life Testing Model

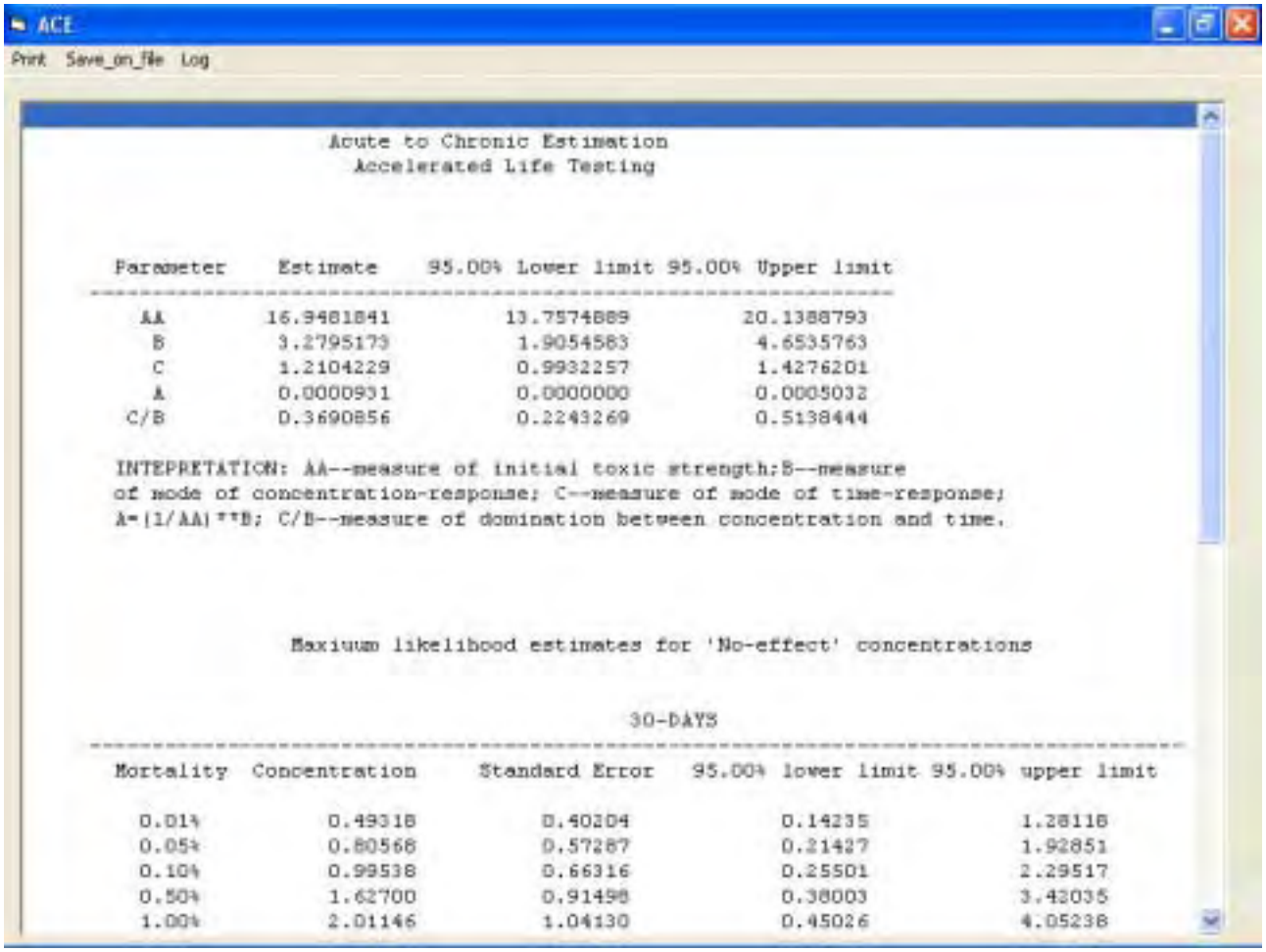
Click on the box **ALT** (Accelerated Life Testing) in the main ACE screen, and analysis of the downloaded acute toxicity data is performed (Fig. 2).

Figure 2 –Accelerated Life Testing (ALT) Screen



Double click on the statistical output screen (left side) in order to obtain the full screen (Fig. 3).

Figure 3 – ALT Full Screen



There are two main parts to the statistical output. The first part contains statistical parameter estimates, along with confidence limits. Interpretation of these parameters follows the estimates. C/B provides an indication of the importance of exposure time (C) versus exposure concentration (B); if equal to one, both are equally important.

The second part of the statistical output is the maximum likelihood estimates of chronic no-effect concentrations. By default, analyses are performed for three different chronic times (30, 60 and 90 days). Within each time period are percent level of chronic mortality (0.01 – 10.0%; 1.0% is recommended for chronic survival with ALT), predicted toxicant concentration associated with each percentage, standard error of the predicted toxicant concentration, and confidence limits (default is 95% confidence limits).

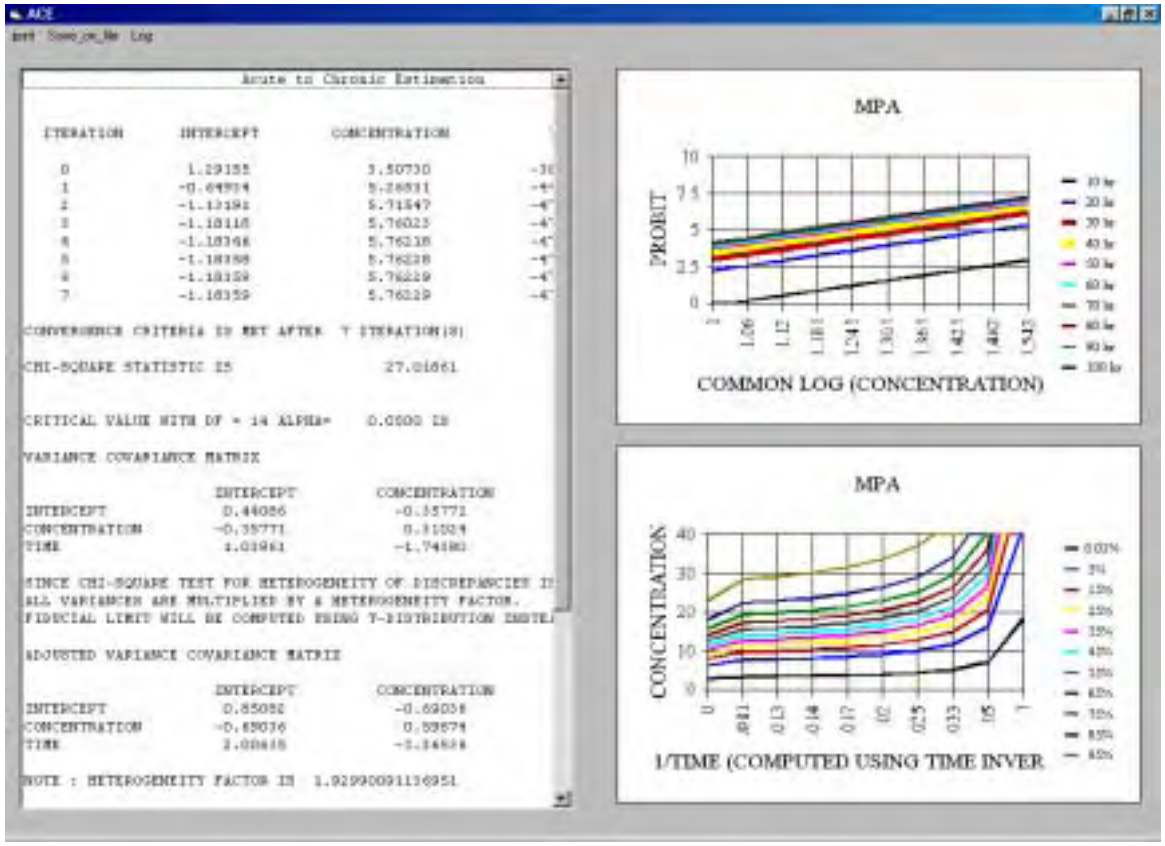
The ALT procedure will function even with a small number of partial responses in the raw acute toxicity data. However, the confidence limits may be large; an error message will appear and the ALT will fail if no partial responses are present in the data.

Additional chronic exposure times and the *alpha* level for confidence limits can be specified (see **Options**).

MPA – Multifactor Probit Analysis Model

Click on the box **MPA** (Multifactor Probit Analysis) in the main ACE screen, and analysis of the downloaded acute toxicity data is performed (Fig. 4).

Figure 4 – Multifactor Probit Analysis (MPA) Screen



Double click on the statistical output screen (left side) in order to obtain the full screen (Fig. 5).

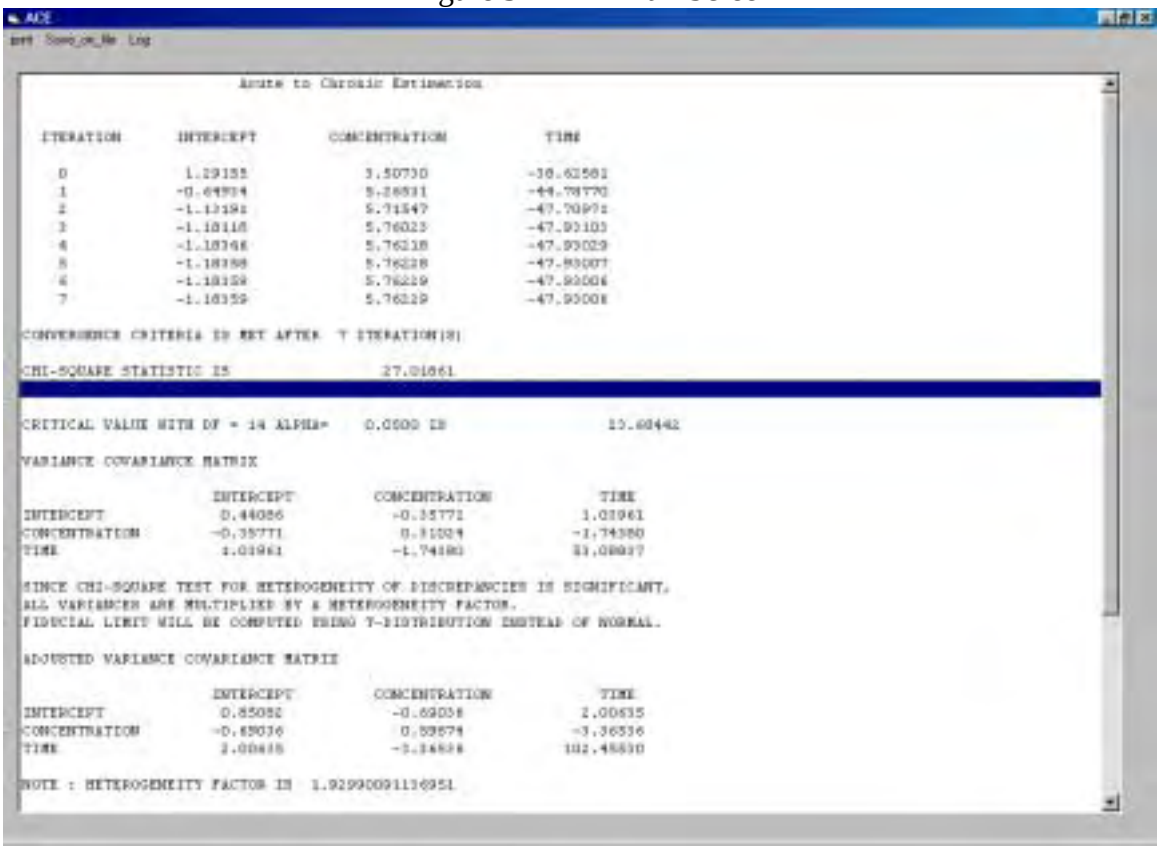
The output provides the number of iterations required to calculate factors for the MPA model, test statistics, variance-covariance matrix, and the predicted chronic no-effect concentrations along with 95% confidence limits.

The MPA includes four different models to choose from that may give different estimates of the MPA functions (see **Options**). The default model is Model 3 in **Options**:

$$Probit_p = \alpha + \beta(Concentration) + \Upsilon/Time$$

Chronic exposure time is specified and the assumption is that slopes change with a constant rate as observation times increase.

Figure 5 – MPA Full Screen



By default, there is one chronic time period (infinity). Within each time period are percent level of mortality (0.01 – 50%; 0.01% is recommended for MPA), predicted toxicant concentration associated with each percentage, and confidence limits (default = 95%). The data fit the model if the chi-square statistic is $\leq$ the critical chi-square value.

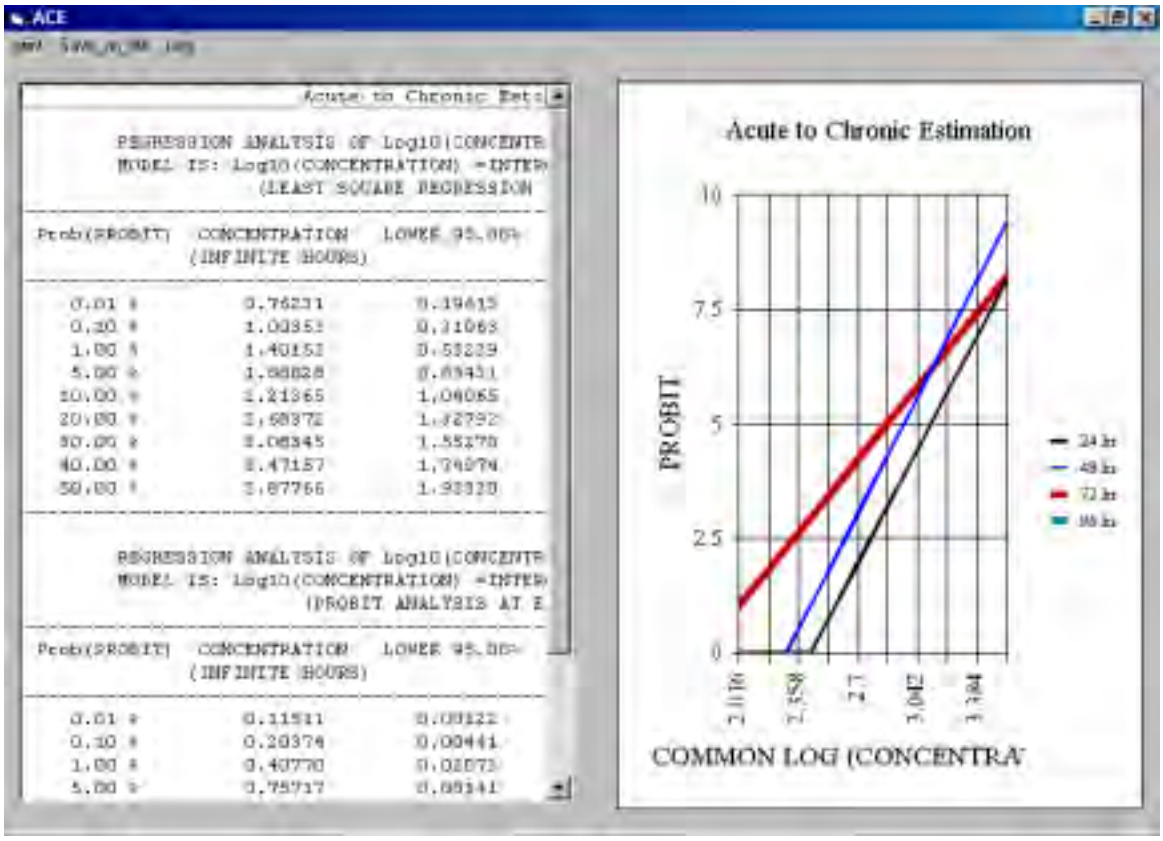
The MPA is the most sensitive to lack of partial mortalities (responses); at least five partial responses between 10 and 90% among all exposure concentrations and times are preferred. An error message will appear and MPA will fail if inadequate partial responses or an insufficient range of partial responses exist.

Additional chronic exposure times and the *alpha* level for confidence limits can be specified (see **Options**).

LRA - Linear Regression Analysis Model

Click on the box **LRA** (Linear Regression Analysis) in the main ACE screen, and analysis of the downloaded acute toxicity data is performed (Fig. 6).

Figure 6 – Linear Regression Analysis (LRA) Screen



Double click on the statistical output screen (left side) in order to obtain the full screen (Fig. 7).

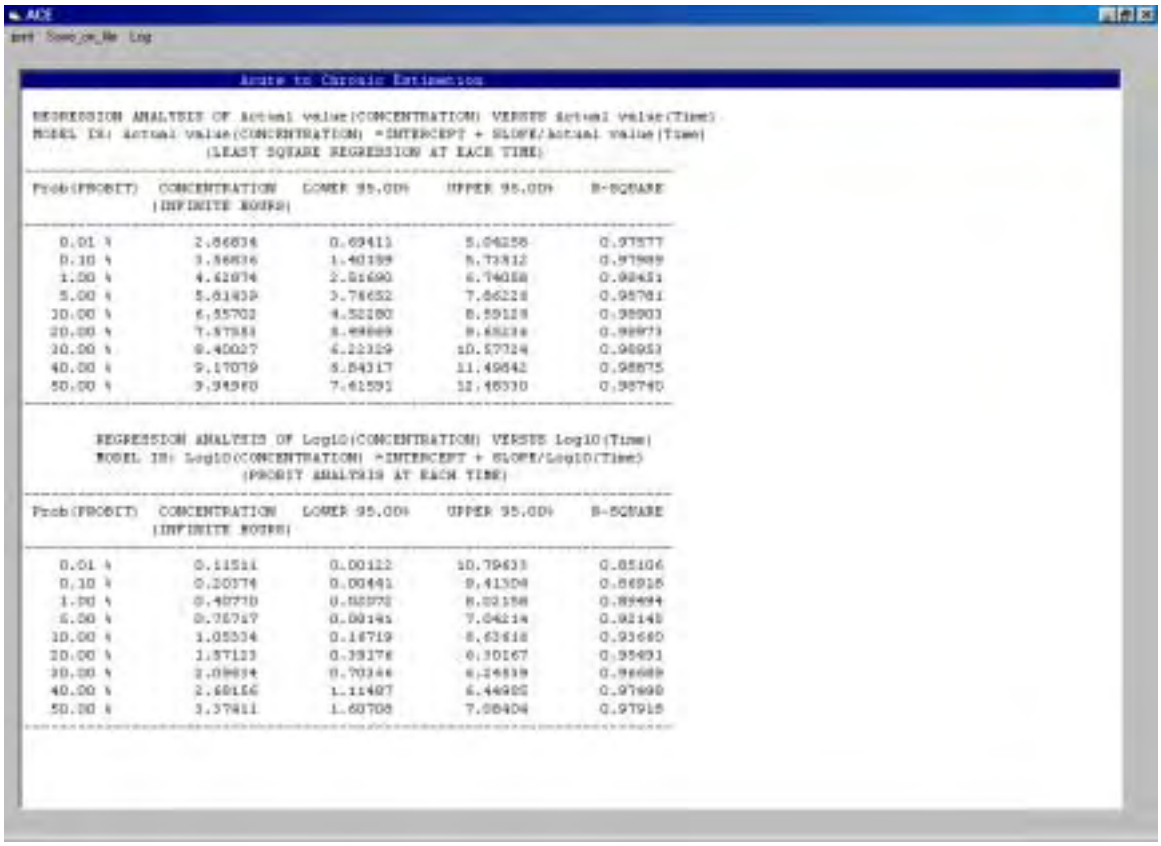
Calculations are based on a two-stage regression analysis, and two analyses will appear; one based on linear regression analysis and another based on probit analysis in stage 1. The following values are given: percent effect (0.01 – 50%; 0.01% is recommended for LRA), estimated chronic no-effect concentration at infinite hours, 95% confidence limits, and r^2. Select the chronic no-effect concentration for 0.01% with the largest r^2 value. Six models are used in the analyses, but only the best stage 1 linear regression and probit analyses (highest r^2) appear in Fig. 7. Click on **Log** to see analyses for all six models.

Note: If percent effects are selected above low percentages (0.01 – 1.0%), abberant values may be apparent when slopes among observation times in stage 1 are very unparallel.

LRA does not require partial responses. If no partial responses are present in the acute toxicity data, LRA uses the highest concentration having 0% (i.e., 0.01%) response at each time period for stage 1, and in stage 2, only the least square analysis is performed.

The only change that can be made for LRA is the *alpha* level for confidence limits (see **Options**); time in hours is set at infinity.

Figure 7 – LRA Full Screen



Options

A number of options are available for controlling the output of each of the ACE models. The options screen is obtained from the main ACE screen. Click **Options** located in the upper left-hand corner of the main ACE window and the following screen will appear (Fig. 8). Once an *alpha* for confidence limits, chronic exposure time, MPA model, and/or statistical output title are changed, click **Save Options**. These changes will remain for present and future analyses. If **Save Options** is not selected, the changes will only remain for the current analysis and then return to default values the next time ACE is used. Click **Restore defaoul options** at the bottom right of the Options window to return to default values.

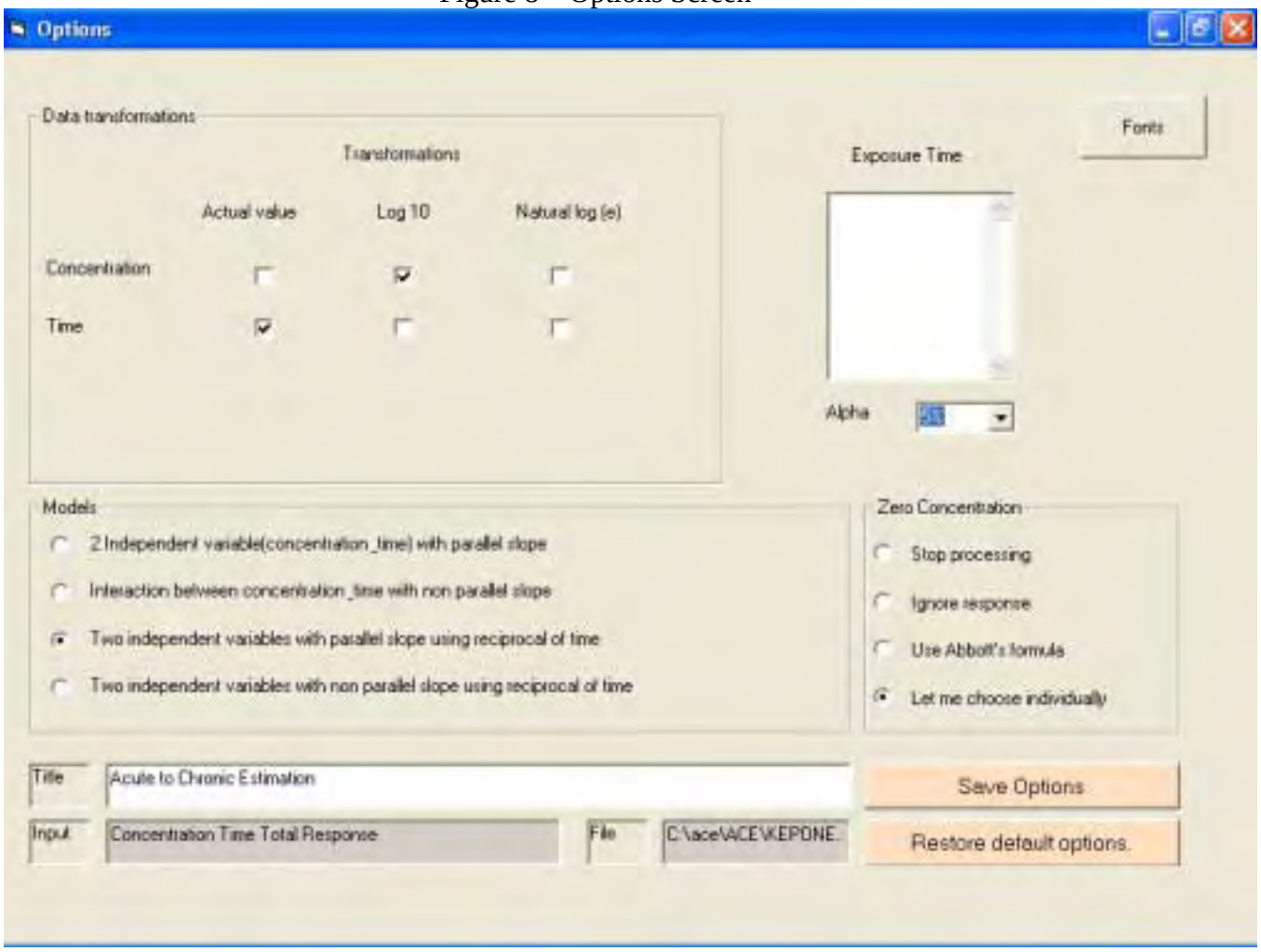
Font

Select **Font** (upper right-hand corner) to change font style of statistical output. Two font styles are presented; fixed font styles should be selected in the left-hand box. The font size may also be changed to fill the data output screen.

Alpha

To change *alpha* levels, click on the arrow associated with **Alpha** located on the upper right side of the **Options** screen; choose the desired *alpha* percent. The *alpha* controls the *t*, z, or chi-square values for producing confidence limits; the *alpha* default value is 5%.

Figure 8 – Options Screen



Exposure Time

In order to change to a different time, go to options window as described previously. To the right is a white box with the header **Exposure Time**. A time change can be accomplished in a number of ways. Type in a number (in hours) in the white box. If a number already exists in the box, write over it or add a number below the existing numbers. No time definition is needed if the number is in terms of hours. However, if one wants to enter days, just type the number of days desired and type "days" after the number and days will be converted to hours by the program. Weeks, months or years can be used as well, by typing in the appropriate time description. Two of these time descriptions (eg., days and months) cannot appear together on the same line. The default for the ALT is 30, 60 and 90 days. The default for MPA is infinite time if the model is based on the reciprocal of time. If the models are not based on the reciprocal of time, a number has to be placed in the **Exposure Time** box in order for the program to calculate NOEC values. The LRA procedure only calculates for infinite time.

Zero Concentration

This section applies only to the MPA and LRA. Abbott's formula (Finney 1978) is used to adjust data if control mortality (zero concentration) exists when probit analysis is performed. The default is **Let me**

choose individually. If control mortality exists, the MPA or LRA will present a message box that allows the user to choose Abbott"s correction. If only one control mortality is present, the message box will appear only once. If control mortality appears more than once, the message box will appear for each one. If **Stop processing** is selected, MPA and LRA will not run if control mortality is present. The **Ignore response** option does not apply Abbott's correction. The **Use Abbott's formula** applies Abbott's correction to all control mortalities.

Title

The title of the statistical output can be changed; click on **Title** and type in a new title. The default title is "Acute to Chronic Estimation".

Selecting MPA Models

The basic Multifactor Probit Analysis equation has a general form in which LC% = Intercept + b_1(Exposure Concentration) + b_2(Time) where b_1 and b_2 are partial regressions for exposure concentration and time, respectively. An additional b_3 [interaction of (exposure concentration)(time)] is added if the slopes among probits are not parallel (see Lee et al. 1995, Mayer et al. 2002).

A number of statistics require evaluation to determine the MPA model of choice. If the chi-square statistic is $\leq$ the critical chi-square value, the data fit the model adequately. Should the other models provide a smaller chi-square statistic, that model is preferred.

To change to one of the other three basic MPA models, exit the MPA program by clicking the **X** in the upper right corner, and then click on **Options** in the upper left-hand corner of the main ACE screen; select **Models** and the four models listed below will appear. The model parameters can be changed to actual values or log values of time and concentration within **Data Transformation** located in the upper left portion of the **Options** screen. This procedure takes much more manipulation to determine the best model. The combination of model choice and actual or log values of concentration and time that gives the lowest chi-square statistics is the best model.

The four models are as follow (1.281 = probit value for 0.01%):

Model 1: Chronic exposure time is specified and equal slopes among observation times are assumed.

Exposure Concentration – Time – Response relationship is defined as:

$$Probit_p = \alpha + \beta(Concentration) + \Upsilon(Time)$$

Chronic no-effect concentration (NOEC) at specified T hours is:

$$NOEC_T = \frac{1.281 - \alpha - \gamma * T}{\beta}$$

Model 2: Chronic exposure time is unknown and equal slopes among observation times are assumed.

Exposure Concentration – Time – Response relationship is defined as:

$$Probit_p = \alpha + \beta(Concentration) + \Upsilon/Time$$

NOEC at infinite time is:

$$NOEC_T = \frac{1.281 - \alpha - \gamma * T}{\beta + \delta + T}$$

Model 3: Chronic exposure time is specified and it is assumed that the slope changes with constant rate as observation times increase.

Exposure Concentration – Time - Response relationship is defined as:

$$Probit_p = \alpha + \beta(Concentration) + \Upsilon(Time) + \delta(Concentration)(Time)$$

NOEC at T hours is:

$$NOEC = \frac{1.281 - \alpha}{\beta}$$

Note: This is the default model in ACE; actual value of time and the log10 of concentration.

Model 4: Chronic exposure time is unknown and and it is assumed that the slope changes with constant rate as observation times increase.

Exposure Concentration – Time – Response relationship is defined as:

$$Probit_p = \alpha + \beta(Concentration) + \Upsilon/Time + \delta\ (Concentration)/(Time)$$

NOEC at infinite time is:

$$NOEC = \frac{1.281 - \alpha}{\beta}$$

Note: Chronic times are necessary for Models 1 and 2; default chronic time is infinity for Models 3 and 4, but additional chronic times may be added.

Estimating Sublethal Effects

Raw data for sublethal endpoints are seldom available under acute exposure conditions for modeling chronic no-effect concentrations. Sublethal endpoints are also difficult to estimate from chronic lethality data. Conservative chronic no-effect concentrations for sublethal endpoints may be estimated by multiplying the predicted NOEC for lethality by 0.2 for growth and other sublethal endpoints and 0.1 for reproductive endpoints. This is based on the analysis of differences among endpoints in chronic toxicity tests (Table 1). However, it must be understood that these estimates of chronic sublethal effects are extremely conservative; note that the median values (that value where 50% of the observations are above or below it) are approximately 1.0 for growth and reproduction and only slightly below 1.0 for "other" sublethal endpoints. In addition, the NOECs for lethality were exactly the same or less than those for weight, length, reproduction, and "other" endpoints 59, 58, 56, and 41% of the time, respectively. Based on the extreme variation of ratios, and the fact that no central tendency exists within the distribution of ratios, **the authors do not recommend using factors to estimate sublethal endpoints at this time**. The data (see table below) are based on hypothesis testing, and using regression analysis to estimate no-effect concentrations for lethal and sublethal endpoints might provide an improved comparison and deserves further investigation.

Univariate analyses for the ratios of growth, reproduction, or other sublethal endpoint chronic no-effect concentrations (NOEC) to that for survival (sublethal NOEC/survival NOEC)[1].

Univariate parameter	Growth		Reproduction	Other[2]
	Weight	Length		
n	46	62	18	22
Mean	0.96	0.90	1.13	0.76
Median	1.0	1.0	1.0	0.6
Range	0.10-4.4	0.16-2.3	0.12-4.5	0.06-2.0
95% CL	0.7-1.2	0.8-1.1	0.6-1.7	0.5-1.0
$\pm$1 SD	0.2-1.8	0.3-1.5	0.1-2.2	0.2-1.3
95th percentile	2.3	2.2	4.5	2.0
Median	1.0	1.0	1.0	0.6
5th Percentile	0.1	0.2	0.1	0.2

[1]Data are from Mayer et al. (1986) and the USEPA Gulf Ecology Division (ORD/NHEERL), Gulf Breeze, FL.
[2]Sublethal endpoints deemed detrimental to survival and/or ability to contribute to population success were cataracts, disease susceptibility, severe fin erosion, severe organ pathology, and spinal curvature.

Additional Model Documentation

Details regarding each model and validation of those models using paired acute and chronic toxicity data are published (Lee et al. 1992, Lee et al. 1995, Mayer 1990, Mayer 1991, Mayer et al. 1992a, Mayer et al. 1992b, Mayer et al. 1994, Mayer et al. 1999, Mayer et al. 2002, Sun et al. 1992, Sun et al. 1994, Sun et al. 1995a, Sun et al. 1995b).

ALT

The ALT procedure uses a Quasi-Newton method to find the maximum likelihood estimates of parameters. Confidence limits for parameters are based on Normal approximations to distributions of the maximum likelihood estimates. The parameter estimates given in Fig. 3 are used in the following model to obtain predicted chronic no-effect concentrations for a particular percent effect and exposure time in days.

$$\text{No-effect concentration} = \text{Exp}[(\ln(-\ln(1-p))-\ln(A) - C*\ln(\text{days}*0.24))/B]$$

A, B, and C are parameter estimates and p is the percent effect, ranging from 0.01 to 10% (see **ALT – Accelerated Life Testing Model**).

MPA

The MPA method uses all time and concentration data simultaneously to produce a multiple regression probit equation to predict chronic no-effect values for specified times.

If the chi-square statistic is $\leq$ the critical chi-square value, a variance-covariance matrix is produced and is necessary to calculate confidence limits. If the chi-square statistic is not $\leq$ the critical chi-square value, the variance-covariance matrix is adjusted by a heterogeneity factor to produce an adjusted variance-covariance matrix. The heterogeneity factor (HF) is given in the statistical output and is equal to the chi-square statistic divided by the degrees of freedom (n – 1 of data used; Finney 1978).

The assumptions of independence may be violated with typical acute toxicity data using MPA. The procedure is appropriate if observations at one time are not the same experimental units at another time. Regardless of the issue of independence, MPA does provide acceptable acute and predicted no-effect chronic concentrations when adequate partial responses are present in the acute data.

LRA

Calculations are based on a two-stage regression analysis. Stage 1 performs two types of analyses. The first type is a simple linear regression at each observation time in which the X axis is log10 concentration and the Y axis is the probit transformation of proportion responding (dead). The second type is a probit analysis at each observation time (Finney 1978). Following these two types of analyses, no-effect concentration values are estimated at different percent response levels. The concentrations are transferred to the stage 2 simple linear regression in which the X axis is the reciprocal of time (1/t) and the Y axis is the concentration at each observation time for a specific percentage value. The equation is:

$c = a + b/t$ where c = chronic no-effect concentration
a = Y intercept
b = regression coefficient
t = time

There are three possible transformations that are made in the stage 2 regression: 1) actual values of concentration and time, 2) log10 of concentration and actual value of time, and 3) log10 of both concentration and time. Thus, six analyses occur due to two types of analyses in stage 1 and three transformations of data in stage 2. As time goes to infinity, the term b/t goes to zero; thus, the concentration at infinite time is the intercept (a), or the chronic no-effect concentration for lethality.

Acknowledgement

This project was sponsored in part by the U.S. Environmental Protection Agency's Offices of Research and Development, Pesticide Programs, Pollution Prevention and Toxics, and Water under Cooperative Agreement CR82827901. Thanks to Vic Camargo for technical support on graphics, and to Debbie Scholes, Mary Adkinson, and Bonnie George for manual preparation. Peer review and beta testing were contributed by M. Anderson, L. Burns, J. Faircloth, T. Linton, R. Pepin, D. Rodier, G. Smith, and W. Waller.

References

Buckler, D.R., A. Witt, Jr.., F.L. Mayer and J. N. Huckins. 1981. Acute and chronic effects of Kepone and Mirex on the fathead minnow. Trans. Am. Fish. Soc. 110:270-280.

Buikema, A.L., Jr., B.R. Nederlehner and J. Cairns, Jr. 1982. Biological monitoring. 4. Toxicity testing. Water Res. 16:239-262.

Compaq Fortran. 1999. Compaq Computer Corporation, Houston, TX.

Crane, M., M.C. Newman, P.F. Chapman and J. Fenlon. 2002. Risk assessment with time to event models. Lewis Publ., Boca Raton, FL. 175 p.

Dixon, P.M. and M.C. Newman. 1991. Analyzing toxicity data using statistical models for time-to-death: An introduction. Pages 207-242, in M.C. Newman and A.W. McIntosh, eds. Metal Ecotoxicology: Concepts and Applications. Lewis Publ., Chelsea, MI.

Finney, D.J. 1978. Statistical methods in biological assay. Griffin, London.

Kenaga, E.E. 1979. Aquatic test organisms and methods useful for assessment of chronic toxicity of chemicals. Pages 101-111, in K.L. Dickson, A.W. Maki and J. Cairns, Jr., eds. Analyzing the Hazard Evaluation Process. American Fisheries Society, Washington, DC.

Kenaga, E.E. 1982. Predictability of chronic toxicity from acute toxicity of chemicals in fish and aquatic invertebrates. Environ. Toxicol. Chem. 1:347-358.

Lee, G., M. Ellersieck and G. Krause. 1992. Multifactor Probit Analysis. Pages 29-61, in F.L. Mayer et al. Statistical Approach to Predicting Chronic Toxicity of Chemicals to Fishes from Acute Toxicity Test Data. National Technical Information Service PB92-169655. U.S. Department of Commerce, Springfield, VA.

Lee, G., M.R. Ellersieck, F.L. Mayer and G. Krause. 1995. Predicting chronic lethality of chemicals to fishes from acute toxicity data: Multifactor probit analysis. Environ. Toxicol. Chem. 14:345-349.

Lipnick, R.L. 1995. Structure-activity relationships. Pages 609-655, in G.M. Rand, ed. Fundamentals of Aquatic Toxicology, 2nd Ed. Taylor & Francis, Washington, DC.

Mayer, F.L., K.S. Mayer and M.R. Ellersieck. 1986. Relationship of survival to other endpoints in chronic toxicity tests with fish. Environ. Toxicol. Chem. 5:737-748.

Mayer, F.L. 1990. Predicting chronic lethality of chemicals to fishes from acute toxicity test data. EPA/600/X-90/147. U.S. Environmental Protection Agency, Gulf Breeze, FL. 15 p.

Mayer, F.L. 1991. Predicting chronic lethality of chemicals to fishes from acute toxicity test data. Pages 56-62 in L.A. Burns, ed. PIRANHA, Pesticide and Industrial Chemical Risk Analysis, Version 2.0. U.S. Environmental Protection Agency, Athens, GA. 134 p. + Appendices.

Mayer, F.L., G.F. Krause, M.R. Ellersieck, and G. Lee. 1992a. Project summary: Statistical approach to predicting chronic toxicity of chemicals to fishes from acute toxicity test data. EPA/600/SR-92/091. U.S. Environmental Protection Agency, Gulf Breeze, FL. 5p.

Mayer, F.L. G.F. Krause, M.R. Ellersieck and G. Lee. 1992b. Statistical approach to predicting chronic toxicity of chemicals to fishes from acute toxicity test data. National Technical Information Service PB92-169644. U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161, 94p. + software.

Mayer, F.L., G.F. Krause, D.R. Buckler, M.R. Ellersieck and G. Lee. 1994. Predicting chronic lethality of chemicals to fishes from acute toxicity data: Concepts and linear regression. Envirol. Toxicol. Chem. 13:671-678.

Mayer, F.L., K. Sun, G. Lee, M.R. Ellersieck, and G.F. Krause. 1999. User guide: Acute to chronic estimation. EPA/600/R-98/152. U.S. Environmental Protection Agency, Washington, DC. 19 p. + software.

Mayer, F.L., M.R. Ellersieck, G.F. Krause, K. Sun, G. Lee, and D.R. Buckler. 2002. Time-concentration-effect models in predicting chronic toxicity from acute toxicity data. Pages 39-67, in M. Crane, M.C.

Newman, P.F. Chapman, and J. Fenlon, eds. Risk Assessment with Time to Event Models. Lewis Publ., Boca Raton, FL.

Microsoft Visual Basic 6.0(SP5). 1987-2000. Microsoft Corporation, United States.

Mount, D.I. and C.E. Stephan. 1967. A method for establishing acceptable limits for fish-malathion and the butoxyethanol ester of 2,4-D. Trans. Am. Fish. Soc. 96:185-193.

Newman, M.C. 1994. Quantitative methods in aquatic ecotoxicology. Lewis Publ., Boca Raton, FL.

Newman, M.C. and M.S. Alpin. 1992. Enhancing toxicity data interpretation and prediction of ecological risk with survival time modeling: An illustration using sodium chloride toxicity to mosquitofish. Aquat. Toxicol. 23:85-96. .

Shirazi, M.A. and L. Lowrie. 1988. Comparative toxicity based on similar asymptotic endpoints. Arch. Environ. Contam. Toxicol. 17:273-280.

Sun, K., G.F. Krause and F. Wright. 1992. Prediction of chemical chronicity using the theory of accelerated life testing: A review. EPA/600/X-92/136. U.S. Environmental Protection Agency, Gulf Breeze, FL. 27 p.

Sun, K., G.F. Krause, F.L. Mayer, M.R. Ellersieck and A.P. Basu. 1994. Predicting chronic toxicity based on the theory of accelerated life testing. EPA/600/R94-058. U.S. Environmental Protection Agency, Gulf Breeze, FL. 33p.

Sun, K., G.F. Krause, F.L. Mayer, M.R. Ellersieck and A.P. Basu. 1995a. Estimation of acute toxicity by fitting a dose-time-response surface. Risk Anal. 15:247-252.

Sun, K., G.F. Krause, F.L. Mayer, M.R. Ellersieck and A.P. Basu. 1995b. Predicting chronic lethality of chemicals to fishes from acute toxicity data: Theory of accelerated life testing. Environ. Toxicol. Chem. 14:1745-1752.

Visual Numeric. 1999. Visual Numerics IMSL Fortran 90 MP Library, Version 4.01 for Microsoft Windows. Visual Numerics, Inc., Houston, TX.

BurrliOZ Help

1 Usage

The BurrliOZ software is designed to estimate the protecting concentrations of chemicals such that a given percentage of species will survive. Before the software can be run, data on the relevant chemicals must be available in ascii files. The software can be run on PCs running Windows NT 4.0, Windows 98 or Windows 2000.

The estimations of the protecting concentrations are computed by fitting a certain distribution to the input data. The distribution, called the Burr III distribution, is that required by the Environment Protection Authority. There are other distributions fitted to the data, the normal and the log-logistic distributions, as these are distributions that the users of this software may be familiar with. However, these two latter distributions are provided only as a reference guide and are not used for the estimation of the protecting concentrations. For further details on the distributions that are fitted and the estimation techniques used see the end of this document.

1.1 Opening a data file for reading

The relevant data file must be opened for reading before the concentration estimation can be done. From the **File** menu select **Open**. A browser window will appear and you can select the desired file. The file should be a readable ascii file, with more than 3 data points listed in it. If this is not the case an error message will appear. In the data file each data point must be on a new line.

1.2 Opening an output file for writing

A file for saving the output must be specified before the protecting concentration can be estimated. From the **File** menu select **Save Output As**. A browser window will appear and you can select an existing file or name a new file in which to save the output. If you select an existing file it will be overwritten.

1.3 Setting percentiles

The percentile is the percentage of species that the estimated concentration should protect. The default is 95%. Accompanied with the percentile is a confidence interval for the estimated concentration. The default for the confidence interval is 50%, which corresponds to no confidence interval being estimated.

To change the percentile and confidence interval select **Set Percentiles** from the **Settings** menu. There are several already specified settings available, which are in the form of ``PC 99 50". This refers to the protecting concentration such that 99% of species survive, with a 50% confidence interval. If **other** is selected from the dialog box by single clicking with the left mouse button, two data entry boxes will appear giving you opportunity to enter a percentile and confidence interval of your choosing. The percentile and the confidence interval should be whole numbers between 0 and 100. If this is not the case an error message will appear. Except for a value of 50%, if the entered confidence interval is less than 80 a warning message will appear. It is conventional to choose a confidence interval of 80%, 85%, 90% or 95%.

1.4 Setting the number of bootstrap samples

If a confidence interval other than 50% is selected, the software must simulate new data in order to estimate the lower confidence limit. The procedure of simulating the data and estimating the confidence

limit from the new data is known as *bootstrapping*. The number of bootstrap samples is the number of data sets simulated to allow for the estimation. The default number is 501, and this should not need to be altered. If changes to this number are required, select **Set number of bootstrap samples** from the **Settings** menu. A sliding bar will appear and a number from 200 to 1000 can be selected. The higher the number, the more accurate the confidence interval estimate will be (but the longer the software will take to run). If the number selected is less than 500, then a warning window will appear when the protecting concentration is estimated, although the estimation procedure will still be allowed to proceed.

1.5 Setting the concentration divisor

The concentration divisor is used to adjust the estimated protecting concentration to give a more conservative level. The choices are *chronic* or *acute*, which give a divisor of 1 or a manually selected value, respectively. Chronic is the default setting. To manually set the divisor select **Set concentration divisor** from the **Settings** menu. If **Acute** is selected in the new dialog box, there will be a choice of dividing by 10 or manually entering an ACR (acute chronic ratio). By default the ACR is 10. Click **OK** when you are happy with your selection.

1.6 Performing the estimation

To compute the protecting concentration for the percentile selected and to estimate the corresponding confidence interval select **Run** from the **File** menu. When this is done a distribution is fitted to the data and the protecting concentration is estimated according to the fitted distribution. A new window will appear with the results displayed and the data and fitted distribution plotted in the window. At the same time as the new window is displayed, the numerical results are saved to the ascii output file that was previously selected.

1.7 Using the results window

The results window shows a plot of the fitted Burr III distribution, along with fits of the normal and log-logistic distributions. Lines are marked on the plot to indicate the estimated protecting concentration (as computed from the Burr III distribution fit).

There are several features of the results window that may be useful in the selection of protecting concentrations that are appropriate for your application. These features are described below.

1.7.1 Highlighting parts of the plot

In the results window a legend appears in the top right hand corner. As the mouse is moved over the legend, a single click on part of the legend will highlight the corresponding part of the plot. More than one part of the plot can be highlighted at one time. The highlighting works as a toggle, so a second click of the left mouse button will revert the corresponding part of the graph back to its original state.

1.7.2 Changing the concentration scale

The scale of the horizontal axis can be on the natural or the log scale. The default is the natural scale. If the log scale is selected, the low end of the horizontal axis will be stretched and the high end will be squashed.

1.7.3 Selecting different percentiles

There is a data entry box in the results window where a different percentile can be typed in. When the

return/enter button is hit the protecting concentration corresponding to the entered percentile is computed. The newly estimated protecting concentration is also marked on the plot. The percentile can only be between 0 and 100.

1.7.4 Selecting concentrations

There is a data entry box in the results window where a protecting concentration can be typed in. When the return/enter button is hit the percentile corresponding to the entered concentration is computed. It is also marked on the plot. The concentration can only be between 0 and the highest concentration shown on the plot.

1.7.5 Zooming

To zoom into part of the plot, click and hold the left mouse button down on the top-left of the desired zoom area. While holding the button down move the mouse to the bottom-right of the desired area and release the button. A single left button mouse click will now complete the zoom. A single right button mouse click will undo the last zoom that was done. This zoom tool may be useful to see the estimated protecting concentration, which can sometimes be too low to see on the original plot.

1.8 Saving results

The results of the estimation procedure are written to the selected output file when **Run** is selected from the **File** menu. To save the results plot simply click the **Save** button on the results window and a Windows metafile will be created and saved to the clipboard. This file can then be pasted into a Word document for future use.

1.9 Printing results

The results plot can be printed by selecting **Print** from the **File** menu in the original BurrliOZ window. A standard Windows printer dialog box will appear and the usual settings can be selected.

1.9.1 Page setup

Some settings for the printing of results can be made in the dialog box that appears by selecting **Page Setup** from the **File** menu in the original BurrliOZ window.

2 Statistical methodology

As outlined at the beginning of this document, the estimations of the protecting concentrations are computed by fitting a certain distribution to the input data. The distribution is called the Burr III distribution, and is the distribution that is required by the Environment Protection Authority. There are other distributions fitted to the data, the normal distribution and the log-logistic distribution, as these are distributions that the users of this software may be familiar with. However, these latter two distributions are provided only as a reference guide and are not used for the estimation of the protecting concentrations.

After the Burr III distribution has been fitted to the data, the protecting concentration (for preserving, for example, 90% of the species) is estimated using the estimated distribution parameters to compute the concentration such that the probability of there being a greater concentration (according to the fitted distribution) is 90%.

After the Burr III distribution has been fitted to the data, the protecting concentration (for preserving, for example, 90% of the species) is estimated using the estimated distribution parameters to compute the concentration such that the probability of there being a greater concentration (according to the fitted distribution) is 90%.

Once the protecting concentration has been computed, an estimate for the lower confidence limit (of a specified percent, say 95%) can be computed. This value can be used as a very conservative (lower) estimate for the protecting concentration.

Further details on how these estimations are done are given below. Some statistical knowledge is required to understand the estimation techniques outlined.

2.1 Fitting the distributions

2.1.1 Burr III fitting

The Burr III distribution is a very flexible 3-parameter distribution, which can provide good approximations to many commonly used distributions such as the log-normal, log-triangular and Weibull. The cumulative distribution function (CDF) for the Burr III distribution is

$$F_{\text{Burr III}}(x) = \frac{1}{\left[1 + \left(\frac{b}{x}\right)^c\right]^k}$$

The 3 parameters of the Burr III distribution, b, c, and k are estimated by maximising the log-likelihood function (which is based on the probability distribution function). This maximisation is performed using the simplex algorithm, an optimisation technique that is not reliant on derivative information.

A complication of the Burr III distribution is that at limits of some of the parameters the Burr III distribution tends to a limiting distribution. As $k \to \infty$ the Burr III distribution tends to the reciprocal Weibull distribution. As $c \to \infty$ the Burr III distribution tends to the reciprocal Pareto distribution. In practical terms, if the Burr III distribution is fitted and k is estimated to be greater than 100, the estimation procedure is carried out again with a reciprocal Weibull distribution fitted. Similarly for the reciprocal Pareto distribution, if c is greater than 80.

Once the parameters are estimated they can be used to compute the CDF of the appropriate distribution (Burr III or one of the limiting distributions), which is plotted with the input dataset in the results window.

2.1.2 Log-logistic fitting

The log-logistic distribution is a commonly used distribution, which happens to be a special case of the Burr III distribution. Because of its possible familiarity with the software user, it has been included as one of the distributions that is fitted to the data and plotted on the results window for comparison against the Burr III fit.

The CDF for the log-logistic distribution is

$$F_{\text{log-logistic}}(x) = \frac{1}{1 + \exp(-[\ln(x) - \mu]/\sigma)}$$

The parameters of the log-logistic distribution, μ and σ, are estimated in the same way as the parameters for the Burr III distribution, by using the simplex algorithm to maximise the log-likelihood function.

2.1.3 Normal fitting

Because the normal (or Gaussian) distribution is a commonly used and well known distribution it has also been included as one of the distributions that is fitted to the data, to allow the user to compare the fit with that of the Burr III distribution. The CDF for the normal distribution is

$$F_{\text{normal}}(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(u-\mu)^2}{2\sigma^2}\right) du$$

For a dataset of n samples, the mean and standard deviation are estimated, respectively, with the maximum likelihood estimators

$$\hat{\mu} = \frac{\sum_{i=1}^{n} x_i}{n}$$

and

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}.$$

As with the log-logistic distribution, the estimated best fit to the data of the normal distribution is also plotted on the results window.

2.2 Estimating the protecting concentration

The protecting concentration is only calculated from the Burr III distribution (or an associated limiting distribution) fit to the data. This is a requirement set out by the Environment Protection Authority. The software user will require the computation of the concentration corresponding to the statement that ``q% of the species should be protected if the concentration of the chemical is less than the estimated concentration". The value of q should be between 0 and 100 and, for the expected use of the BurrliOZ software, will be somewhere close to 100, such as 80, 85, 90 or 95.

For a given value for q, the protecting concentration is estimated from the Burr III distribution fit as

$$PC(q) = \frac{b}{[(1/(1-q))^{1/k} - 1]^{1/c}}.$$

If the limiting distribution of reciprocal Weibull is used then the protecting concentration is estimated as

$$PC(q) = (-\alpha/\ln(1-q))^{1/\beta}$$

where α and β are the two parameters of the reciprocal Weibull distribution that were estimated in the fitting step. Similarly, if the reciprocal Pareto distribution has been necessarily fitted, the protecting concentration is estimated as

$$PC(q) = x_0 (1-q)^{1/\theta}$$

where x_0 and θ are the two parameters of the reciprocal Pareto distribution that were estimated in the fitting step.

2.3 Estimating a confidence interval for the protecting concentration

Unlike the estimation of the protecting concentration (PC), there is no theoretically derived equation for estimating the lower bound of a confidence interval (CI) about the protecting concentration estimate. Instead, a technique known as bootstrapping is used to estimate the lower bound of the CI.

To perform the bootstrapping, a new dataset of the same size as the original dataset is created by selecting values from the original set at random with replacement. With the new dataset, the PC is estimated using the steps outlined above. This process is repeated many times (the default being 501). This gives a large set of estimates for the PC. In essence, this is a representation of the distribution of the PC. The lower bound of a 90% confidence interval (for example) for the PC can then be estimated by ordering all the PC values and selecting the value that is 5% in from the lowest value.

It should be noted that the estimated lower bound to the CI is based on a random sampling method and will not be the same when the bootstrap is repeated.

File translated from TEX by TTH, version 2.60.
On 25 May 2000, 15:16.

Environmetrics

Home

Track Record

Skills and Research

- Hybrid Climate Modelling: Physical Statistical Models

Current Projects

Team

BurrliOZ

A Flexible Approach to Species Protection

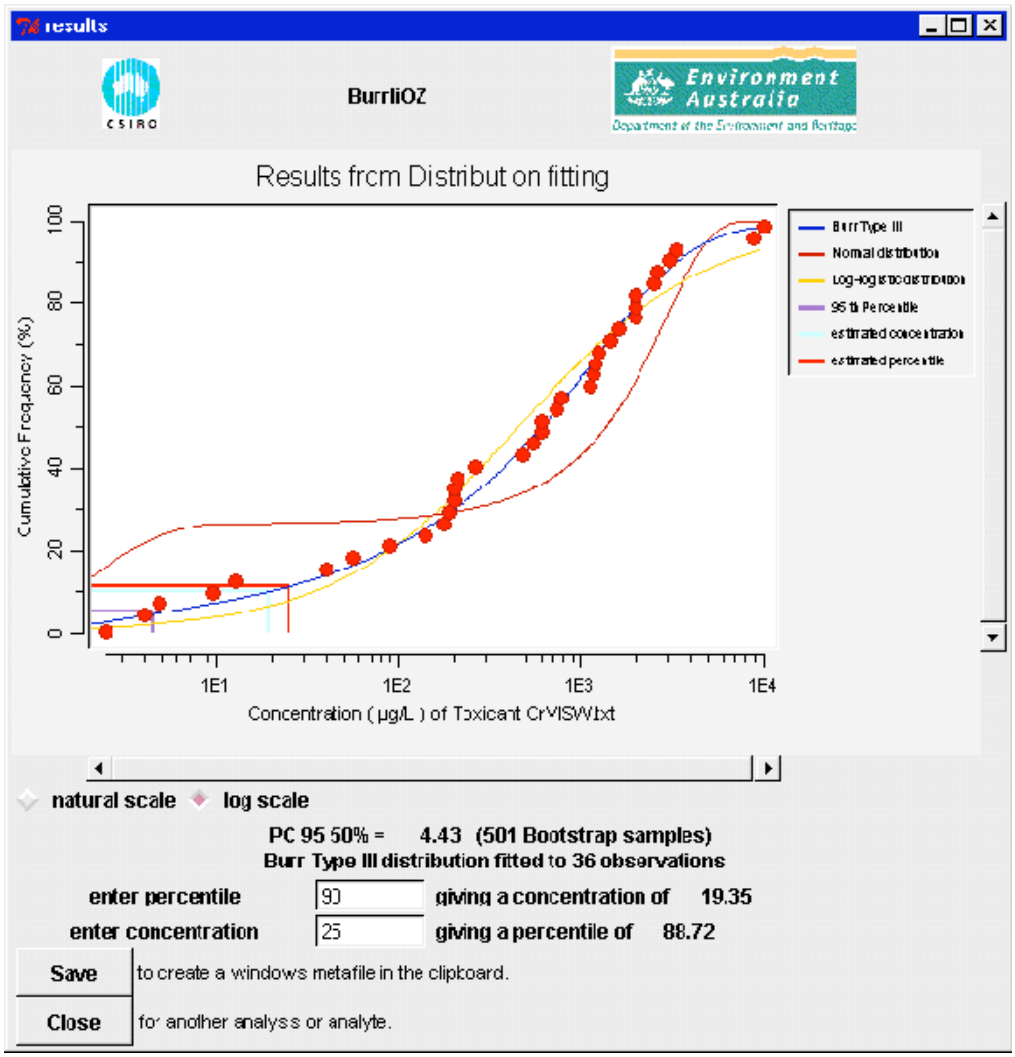
Environmental Managers responsible for implementing the Australian and New Zealand Water Quality Guidelines for Fresh and Marine Waters need to generate 'trigger values' (ie the maximum concentration of a chemical that should permit the integrity and function of aquatic environments to be maintained) for local conditions within Australia. To do this, they will utilise toxicant data and a statistical software package, BurrliOZ, developed by the CSIRO Environmetrics Group for Environment Australia. Another software package that calculates trigger values using the Aldenberg and Slob (1993) approach exists, however this has been shown to be a special case of the approach implemented in BurrliOZ (Shao, 2000). BurrliOZ uses a flexible family of distributions, the Burr Type III, to estimate the concentrations of chemicals such that a given percentage of species will survive.

Environment Australia
Department of the Environment and Heritage

This project makes available to the public, free of charge and subject to certain restrictions, a new software packages including both the 'Web' and a CD-ROM suitable for delivery with the Guidelines document. This software and delivery format addresses concerns raised during the 1999 public comment period. Download Burrlioz software.

The work is expected to facilitate approval of the final Water Quality Guidelines by the Australian and New Zealand Environment and Conservation Council (ANZECC) and the Agriculture and Resource Management Council of Australia and New Zealand (ARMCANZ) Ministers and also accelerate effective implementation of the Water Quality Guidelines. This work represents a significant advance in the methods used to derive water quality guidelines and it should have international implications and uses.

A screen shot of what BurrliOZ looks like



The Method

The protecting concentrations are estimated by fitting the Burr Type III distribution to the No Observed Effect Concentration (NOEC) data, collected for a range of species. This distribution is required by the Environment Protection Authority. Other distributions are fitted to the data, including the log-normal and log-logistic as these are familiar to environmental managers. However, they are provided only as a reference and are not used for the estimation of protecting concentrations.

The Burr III distribution is a very flexible three-parameter distribution, which can provide good approximations to many commonly used distributions such as the log-normal, log-triangular and Weibull. The cumulative distribution function for the Burr III distribution is

$$F(x) = \frac{1}{\left[1 + \left(\frac{b}{x}\right)^{c}\right]^{k}}$$

The three-parameters of the Burr III distribution, *b*, *c*, and *k* are estimated by maximum likelihood using the Nelder-Mead simplex algorithm, a derivative free optimisation technique.

A feature of the Burr Type III distribution is that as some of the parameters tend to limiting values the Burr Type III distribution tends to one of a set of limiting distribution (Shao, 2000). For example, as $k \to \infty$ the Burr III distribution tends to the reciprocal Weibull distribution. As $c \to \infty$ the Burr III distribution tends to the reciprocal Pareto distribution. In practice, if *k* is estimated to be greater than 100 in a fit of the Burr distribution, then the parameter estimation is repeated, a reciprocal Weibull is fitted. Similarly if *c* is estimated to be greater than 80 then the reciprocal Pareto distribution is fitted.

Estimating the protecting concentration

The protecting concentration, PC(q), is calculated from the Burr Type III distribution, or an associated limiting distribution. The user requires the concentration corresponding to the statement that ``q% of the species should be protected if the concentration of the chemical is less than the estimated protecting concentration". Thus, for a given value for q, the protecting concentration is estimated from the Burr III distribution fit as

$$PC(q) = \frac{b}{\left[\left(\frac{1}{1-q}\right)^{\frac{1}{k}} - 1\right]^{\frac{1}{c}}}$$

Typical values for q are 80, 85, 90 or 95.

Estimating a confidence interval for the protecting concentration

Unlike the estimation of the protecting concentration, there is no theoretically derived equation for estimating the lower bound of a confidence interval (CI) about the protecting concentration etimate, though Shao (1998) has shown that a delta method approximation works sometimes, particularly for large samples. Instead, a technique known as bootstrapping is used to estimate the lower bound of the CI. Bootstrapping is a standard statistical approach in situations where theoretical results are difficult to obtain, or require unrealistic assumptions (Efron and Tibshirani, 1993).

To perform the bootstrapping, a new dataset of the same size as the original dataset is created by selecting values from the original set at random, but with replacement. The PC(q) is estimated from this new dataset as above. This process is repeated many times. This gives a large set of estimates for the PC(q) which, in essence, is a representation of the distribution of the PC(q). The lower bound of a 90% confidence interval (for example) for the PC(q) can then be estimated by ordering all the PC(q) values and selecting the value that is ranked at 5%.

It should be noted that the estimated lower bound to the CI is based on a random sampling method and will not be exactly the same if the bootstrap procedure is repeated.

References:

Aldenberg, T. and Slob, W. (1993). Confidence limits for hazardous concentrations based on logistically distributed NOEC toxicity data. *Ecotoxicology and Environmental Safety*, **25**, 48-63

Efron, B. and Tibshirani, R.J. (1993). *An introduction to the Bootstrap*. New York: Chapman & Hall.

Shao, Q. (1998). *Statistical Review and Assessment of Water Quality Guidelines*, CSIRO Mathematical and Information Sciences Report No CMIS98/21

Shao, Q. (2000). Estimation for hazardous concentrations based on NOEC toxicity data: an alternative approach. (accepted by *Environmetrics*)

Contact Details: burrlioz@cmis.csiro.au

Download Burrlioz software.

To top

last updated November 07, 2002 11:52 AM
Bert.deBoer@cmis.csiro.au

Canadian Environmental Modelling Centre

Level I Model

Version 3.00 - September 2004

New in version 3.00!

A Level I simulation is of the equilibrium distribution of a fixed quantity of conserved (ie. non-reacting) chemical, in a closed environment at equilibrium, with no degrading reactions, no advective processes, and no intermedia transport processes (e.g. no wet deposition, or sedimentation). The medium receiving the emission is unimportant because the chemical is assumed to become instantaneously distributed to an equilibrium condition.

Physical-chemical properties are used to quantify a chemical's behaviour in an evaluative environment. Three types of chemicals are treated in this model: chemicals that partition into all media (Type 1), involatile chemicals (Type 2), and chemicals with zero, or near-zero, solubility (Type 3). The Level I Model assumes a simple, evaluative, closed environment with user-defined volumes and densities for the following homogeneous environmental media (or compartments): air, water, soil, sediment, suspended sediment, fish and aerosols.

This model is useful for establishing the general features of a new or existing chemical's behaviour. A Level I calculation gives the general impression of the likely media into which a chemical will tend to partition and an indication of relative concentrations in each medium. The results of changes in chemical and environmental properties may be explored.

Features of the Level I Program:
Provides a database of chemicals and chemical properties.
Permits temporary additions/changes of chemicals and their properties to a simulation.
Permits permanent additions, changes and deletions of chemicals and their properties to the chemical database.
Supplies default values for all input fields which may be easily changed. These values are regarded as typical, as discussed in the text referred to earlier.
Provides context-sensitive Help.
Displays and prints the Level I model calculations, as performed by the program.
Allows the printing of simulation tables and the summary diagram.
Allows the program results to be saved as a comma separated value (csv) file.

This program was based on the following publication:

Mackay, D. 2001. "Multimedia Environmental Models: The Fugacity Approach - Second Edition", Lewis Publishers, Boca Raton, pp. 1-261.

Other related publications:

Mackay, D., Paterson, S., Kicsi, G., Di Guardo, A., Cowan, C.E. 1996. Assessing the Fate of New and Existing Chemicals: A Five Stage Process. Environ. Toxicol. Chem. 15: 1618-1626.

Mackay, D., Paterson, S., Di Guardo, A., Cowan, E.C. 1996. Evaluating the Environmental Fate of a Variety of Types of Chemicals Using the EQC Model. Environ. Toxicol. Chem. 15: 1627-1637.

Mackay, D., Paterson, S., Kicsi, G., Cowan, E.C., Di Guardo, A., Kane, D.M. 1996. Assessment of Chemical Fate in the Environment Using Evaluative, Regional and Local-Scale Models: Illustrative Application to Chlorobenzene and Linear Alkylbenzene Sulfonates. Environ. Toxicol. Chem. 15: 1638-1648.

The required input data are:
Chemical Properties:

- chemical name
- molecular mass
- data temperature
- Type 1 chemicals

- water solubility
- vapour pressure
- log Kow
- melting point

- Type 2 and 3 chemicals

- partition coefficients

Environmental Properties:

- volumes for all 7 media
- densities for all 7 media
- organic carbon content (soil, sediment & suspended sediment only)
- fish lipid content (Type I chemicals only)

Emissions:

- chemical amount

Model Output:

- partition coefficients (Type 1)
- Z values
- fugacity of the system
- concentrations and amounts for each compartment
- a summary diagram

This program is only available in compiled form. A "readme.txt" file with more detailed technical information is included in the zipped file.

Minimum sytem requirements:
Pentium-75MHz with 8 Mb of RAM running Windows 95, 98, or XP. On some systems it may be necessary to adjust your screen resolution.

The Level I Model Version 2.11, released August 1999 continues to be available.

For non-Windows users the BASIC, evaluative, Level I, II and III fugacity models are available.

Please read the **LEVEL I SOFTWARE LICENSE** before downloading the software. Use of the software constitutes your agreement to abide by the terms and conditions set out in the license agreement.

Register to Download

CEMC Home | General Information | Models | Graduate Studies | Contact

Last updated September 24, 2004.

Canadian Environmental Modelling Centre

Level I Software License

Version 3.00 - September 2004

The Level I software program is provided to interested parties at no cost. We do request that you provide us with registration information at the time of download from our website for our own information and statistical purposes.

The use of the Level I software program is governed by the following legal agreement. The purpose of the agreement is to ensure that all users are treated equitably, and we are not disadvantaged in any way by providing the software. All use of this software is conditional upon your compliance with the license terms which follow. If you do not agree to the terms of this license agreement, or do not comply with the terms and conditions of this agreement, you are not permitted to use this software and are required to remove the Level I software program from your computer system, and destroy all copies of the software. The use of the Level I software program constitutes your agreement to abide by the terms and conditions set out in this document.

THE Level I SOFTWARE PROGRAM, HEREIN CALLED THE "SOFTWARE", IS OWNED BY TRENT UNIVERSITY AND IS PROTECTED BY CANADIAN COPYRIGHT LAWS. UPON YOUR AGREEMENT TO AND COMPLIANCE WITH THE TERMS OF THIS LICENSE AGREEMENT, TRENT UNIVERSITY GRANTS YOU, HEREIN CALLED THE "LICENSEE", THE FOLLOWING NON-TRANSFERRABLE, NON-EXCLUSIVE RIGHTS OF USE. TRENT UNIVERSITY HAS THE RIGHT TO TERMINATE THIS AGREEMENT IF THE "LICENSEE" FAILS TO COMPLY WITH ANY TERM OR CONDITION OF THIS AGREEMENT. NO TITLE TO THE INTELLECTUAL PROPERTY IN THE "SOFTWARE" IS TRANSFERRED TO YOU. THE "LICENSEE" DOES NOT ACQUIRE ANY RIGHTS TO THE "SOFTWARE" EXCEPT AS EXPRESSLY SET FORTH IN THIS LICENSE.

GRANT OF LICENSE REGARDING THE Level I PROGRAM

TRENT UNIVERSITY GRANTS THE "LICENSEE" THE FOLLOWING RIGHTS REGARDING THE USE OF THE "SOFTWARE":

1) USE OF THE "SOFTWARE" FOR THE "LICENSEE'S" PERSONAL OR BUSINESS PURPOSES.

2) COPYING THE "SOFTWARE"
i) THE "LICENSEE" MAY NOT MAKE COPIES OF THE "SOFTWARE" OTHER THAN THOSE GRANTED BY LAW FOR ARCHIVAL OR BACKUP PURPOSES.
ii) THE "SOFTWARE" MAY BE TRANSFERRED TO THE HARD DISK OF ANY COMPUTER, OR NETWORK OF COMPUTERS, BELONGING TO THE "LICENSEE".

RESTRICTIONS REGARDING THE Level I PROGRAM

1) THE "LICENSEE" MAY NOT REVERSE ENGINEER, DECOMPILE, DISASSEMBLE, MODIFY, TRANSLATE, OR ALTER THE SOFTWARE AND/OR THE ASSOCIATED FILES, OR IN ANY MANNER SUPPORT OR CAUSE SUCH TO OCCUR.

2) THE "LICENSEE" MAY NOT DISTRIBUTE, SUBLICENSE, LEASE, SELL, RENT OR OTHERWISE TRANSFER THE "SOFTWARE", OR ANY MODIFICATION OR DERIVATIVE THEREOF, TO ANY OTHER INDIVIDUAL OR GROUP FOR ANY REASON. (Additional copies can be obtained directly from this web site.)

LIMITATIONS OF LIABILITY AND DISCLAIMER OF WARRANTY

THERE ARE NO WARRANTY RIGHTS GRANTED TO YOU, THE "LICENSEE", REGARDING THE "SOFTWARE". THE "SOFTWARE" AND ACCOMPANYING WRITTEN MATERIALS ARE SUPPLIED TO THE "LICENSEE" "AS IS" WITHOUT WARRANTY OF ANY KIND. TRENT UNIVERSITY DOES NOT GUARANTEE, WARRANT, OR MAKE ANY REPRESENTATIONS, EITHER EXPRESSED OR IMPLIED, REGARDING THE USE, OR THE RESULTS OF THE USE OF THE THE "SOFTWARE" OR THE Level I WRITTEN MATERIALS WITH REGARDS TO RELIABILITY, CURRENTNESS, ACCURACY, CORRECTNESS, OR OTHERWISE. THE "LICENSEE" ASSUMES THE ENTIRE RISK AS TO THE RESULTS AND PERFORMANCE OF THE "SOFTWARE".

TRENT UNIVERSITY SHALL NOT BE LIABLE UNDER ANY CIRCUMSTANCES, FOR ANY DAMAGES WHATSOEVER, ARISING OUT OF THE USE, OR THE INABILITY TO USE, THE "SOFTWARE", EVEN IF TRENT UNIVERSITY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

Model Description

CEMC Home | General Information | Models | Graduate Studies | Contact

Last updated September 24, 2004.

Canadian Environmental Modelling Centre

Level I Software Update

Version 3.00 - September 2004

Since the Version 2.11 (August, 1999) release:

- A database of Environmental Properties was added.
- General layout and functionality were updated to be consistent with recent CEMC model releases.

Model Description

CEMC Home | General Information | Models | Graduate Studies | Contact

Last updated September 23, 2004.

Canadian Environmental Modelling Centre

Level II Model

Version 3.00 - September 2005

New in version 3.00!

A Level II simulation describes a situation in which a chemical is continously discharged at a constant rate and achieves a steady-state and equilibrium condition at which the input and output rates are equal. Degrading reactions and advective processes are the loss or output processes treated. Intermedia transport processes (e.g. no wet deposition, or sedimentation) are not quantified. The medium receiving the emission is unimportant because the chemical is assumed to become instantaneously distributed to an equilibrium condition.

Physical-chemical properties are used to quantify a chemical's behaviour in an evaluative environment. Three types of chemicals are treated in this model: chemicals that partition into all media (Type 1), involatile chemicals (Type 2), and chemicals with zero, or near-zero, solubility (Type 3). The Level II model assumes a simple, evaluative environment with user-defined volumes and densities for the following homogeneous environmental media (or compartments): air, water, soil, sediment, suspended particles, fish and aerosols.

This model is useful for establishing the general features of a new or existing chemical's behaviour. A Level II calculation gives an indication of the likely media into which a chemical will tend to partition and an indication of relative concentrations in each medium. The distribution between media is the same as in Level I. The results of changes in chemical and environmental properties may be explored.

Three persistences are calculated, an overall value, T_O, and individual persistences attributable to reaction only, T_R, and advection only, T_A. Note that $1/T_O$ equals the sum of $1/T_R$ and $1/T_A$.

Consideration of advection and reaction rates allows for the calculation of chemical persistence. It provides a first estimate of overall environmental persistence, which is a critical property of the chemical. It also shows which loss processes are likely to be most important. A fast reaction or short half-life may not be significant if relatively little of the chemical is subject to this reaction by virtue of its partitioning. The potential for the chemical to be subject to long-range atmospheric transport is also indicated by the magnitude of the air advection loss. The global chemical persistence is best indicated by the reaction persistence, whereas the local persistence is indicated by the overall persistence.

Note that in this version, reaction half-lives are requested for all 7 media. In previous versions reactions in only 4 media were treated. The advective residence time selected for air also applies to aerosols and the residence time for water applies to suspended particles and fish. The advective residence time of aerosols, suspended particles and fish cannot be specified independently of the air and water residence times.

A Level II calculation is more realistic than a Level I calculation but requires additional information.

Features of the Level II Program:
Provides a database of chemicals and chemical properties.
Permits temporary additions/changes of chemicals and their properties to a simulation.

Permits permanent additions, changes and deletions of chemicals and their properties to the chemical database.
Supplies default values for all input fields which may be easily changed. These values match those in the EQC model.
Provides context-sensitive Help.
Displays and prints the Level II model calculations, as performed by the program.
Allows the printing of simulation tables and the summary diagram and charts.
Allows the program results to be saved as a comma separated value (csv) file readable by most spreadsheet software.

This program was based on the following publication:

Mackay, D. 2001. "Multimedia Environmental Models: The Fugacity Approach - Second Edition", Lewis Publishers, Boca Raton.

The required input data are:
Chemical Properties:

- chemical name
- molar mass
- data temperature
- reaction half-life estimates for

- air
- aerosols
- water
- suspended particles
- aquatic biota
- soil
- sediment

- Type 1 chemicals

- water solubility
- vapour pressure
- log Kow
- melting point

- Type 2 and 3 chemicals

- partition coefficients

Environmental Properties:

- volumes for all media
- densities for all media
- organic carbon content (soil, sediment, and suspended particles only)
- fish lipid content
- advective flow residence times for air (including aerosols), and water (including suspended particles and aquatic biota)
- advective flow residence time for sediment burial

Emissions:

- chemical input rate
- inflowing concentrations in air and water

Model Output:

- partition coefficients (Type 1)
- Z values
- fugacity of the system
- D values
- reaction and advection loss rates
- residence times or persistences (overall, reaction, and advection)
- concentrations and amounts for each compartment
- a summary diagram and charts

Minimum system requirements are an IBM-compatible PC running Windows 98 or XP.
This model will not run under Windows NT or 2000.

This program is only available in compiled form. A "readme.txt" file with more detailed technical information is included in the zipped file.

The Level II Model Version 2.17, released September 1999 continues to be available.

For non-Windows users the BASIC, evaluative, Level I, II and III fugacity models are available.

Please read the **LEVEL II SOFTWARE LICENSE** before downloading the software. Use of the software constitutes your agreement to abide by the terms and conditions set out in the license agreement.

Register to Download

CEMC Home | General Information | Models | Graduate Studies | Contact

Last updated October 26, 2005.

Canadian Environmental Modelling Centre

Level II Software License

Version 3.00 - September 2005

The Level II software program is provided to interested parties at no cost. We do request that you provide us with registration information at the time of download from our website for our own information and statistical purposes.

The use of the Level II software program is governed by the following legal agreement. The purpose of the agreement is to ensure that all users are treated equitably, and we are not disadvantaged in any way by providing the software. All use of this software is conditional upon your compliance with the license terms which follow. If you do not agree to the terms of this license agreement, or do not comply with the terms and conditions of this agreement, you are not permitted to use this software and are required to remove the Level II software program from your computer system, and destroy all copies of the software. The use of the Level II software program constitutes your agreement to abide by the terms and conditions set out in this document.

THE LEVEL II SOFTWARE PROGRAM, HEREIN CALLED THE "SOFTWARE", IS OWNED BY TRENT UNIVERSITY AND IS PROTECTED BY CANADIAN COPYRIGHT LAWS. UPON YOUR AGREEMENT TO AND COMPLIANCE WITH THE TERMS OF THIS LICENSE AGREEMENT, TRENT UNIVERSITY GRANTS YOU, HEREIN CALLED THE "LICENSEE", THE FOLLOWING NON-TRANSFERRABLE, NON-EXCLUSIVE RIGHTS OF USE. TRENT UNIVERSITY HAS THE RIGHT TO TERMINATE THIS AGREEMENT IF THE "LICENSEE" FAILS TO COMPLY WITH ANY TERM OR CONDITION OF THIS AGREEMENT. NO TITLE TO THE INTELLECTUAL PROPERTY IN THE "SOFTWARE" IS TRANSFERRED TO YOU. THE "LICENSEE" DOES NOT ACQUIRE ANY RIGHTS TO THE "SOFTWARE" EXCEPT AS EXPRESSLY SET FORTH IN THIS LICENSE.

GRANT OF LICENSE REGARDING THE LEVEL II PROGRAM

TRENT UNIVERSITY GRANTS THE "LICENSEE" THE FOLLOWING RIGHTS REGARDING THE USE OF THE "SOFTWARE":

1) USE OF THE "SOFTWARE" FOR THE "LICENSEE'S" PERSONAL OR BUSINESS PURPOSES.

2) COPYING THE "SOFTWARE"
i) THE "LICENSEE" MAY NOT MAKE COPIES OF THE "SOFTWARE" OTHER THAN THOSE GRANTED BY LAW FOR ARCHIVAL OR BACKUP PURPOSES.
ii) THE "SOFTWARE" MAY BE TRANSFERRED TO THE HARD DISK OF ANY COMPUTER, OR NETWORK OF COMPUTERS, BELONGING TO THE "LICENSEE".

RESTRICTIONS REGARDING THE LEVEL II PROGRAM

1) THE "LICENSEE" MAY NOT REVERSE ENGINEER, DECOMPILE, DISASSEMBLE, MODIFY, TRANSLATE, OR ALTER THE SOFTWARE AND/OR THE ASSOCIATED FILES, OR IN ANY MANNER SUPPORT OR CAUSE SUCH TO OCCUR.

2) THE "LICENSEE" MAY NOT DISTRIBUTE, SUBLICENSE, LEASE, SELL, RENT OR OTHERWISE TRANSFER THE "SOFTWARE", OR ANY MODIFICATION OR DERIVATIVE THEREOF, TO ANY OTHER INDIVIDUAL OR GROUP FOR ANY REASON. (Additional copies can be obtained directly from this web site.)

LIMITATIONS OF LIABILITY AND DISCLAIMER OF WARRANTY

THERE ARE NO WARRANTY RIGHTS GRANTED TO YOU, THE "LICENSEE", REGARDING THE "SOFTWARE". THE "SOFTWARE" AND ACCOMPANYING WRITTEN MATERIALS ARE SUPPLIED TO THE "LICENSEE" "AS IS" WITHOUT WARRANTY OF ANY KIND. TRENT UNIVERSITY DOES NOT GUARANTEE, WARRANT, OR MAKE ANY REPRESENTATIONS, EITHER EXPRESSED OR IMPLIED, REGARDING THE USE, OR THE RESULTS OF THE USE OF THE THE "SOFTWARE" OR THE LEVEL II WRITTEN MATERIALS WITH REGARDS TO RELIABILITY, CURRENTNESS, ACCURACY, CORRECTNESS, OR OTHERWISE. THE "LICENSEE" ASSUMES THE ENTIRE RISK AS TO THE RESULTS AND PERFORMANCE OF THE "SOFTWARE".

TRENT UNIVERSITY SHALL NOT BE LIABLE UNDER ANY CIRCUMSTANCES, FOR ANY DAMAGES WHATSOEVER, ARISING OUT OF THE USE, OR THE INABILITY TO USE, THE "SOFTWARE", EVEN IF TRENT UNIVERSITY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

Model Description

CEMC Home | General Information | Models | Graduate Studies | Contact

Last updated October 25, 2005.

Canadian Environmental Modelling Centre

Level II Software Update

Version 3.00 - September 2005

Since the Version 2.17 (September 1999) release:

Simulation ID

- There is now an input box for additional notes which is then inputted directly into the print simulation section and will appear on the final printout.

Chemical Properties

- There is a change to the half-life checkboxes. When the half-life checkbox is ticked so that the half-life is considered to be negligible and then the user enters a new half-life into the input box the check box is unchecked and the new value is stored.
- The Henry's Law Constant display has been removed from this form and is now only viewable in the results form. This has no affect the user's results.

Environmental Properties

- Addition of an environmental database. This database functions in the same manner that the chemical database functions.
- A water volume of 0 m^3 is now allowed. When this value is entered it will automatically change the volumes for sediment, fish, and suspended particles to zero m^3.

Emissions

- There are no changes to this section.

Chemical Parameters

- Same information is being displayed, it is organized differently.

Environmental Parameters

- Emissions and Inflows are now displayed on the Results form.

Results

- Unit conversions are displayed as options as opposed to buttons.
- Modified grids to match format of Level III.

- Concentration now displayed in alternate units.

Diagram

- The same information is being displayed, but appearance has been updated.

Charts

- Addition of charts to display Concentration (g/m^3), Amount (kg), and Relative Amount (%)

Error Checking

- Error checking for single corrected.

Calculations

- Fixed print out of calculations, it now correctly prints same information that is displayed

Print

- The option for printing the date and time when printing Tables has been corrected so that it works.
- Note: All corrections to units and calculations in above sections have carried over to the printout.

Save to File

- Now saves file as .csv format instead of a .txt format.

About

- The New Features Button.

Model Description

CEMC Home | General Information | Models | Graduate Studies | Contact

Last updated October 26, 2005.

Canadian Environmental Modelling Centre

Level III Model

Version 2.80 - May 2004

New in version 2.80!

A Level III simulation describes a situation which is one step more complex and realistic than the Level II model. Like the Level II model, chemical is continuously discharged at a constant rate and achieves a steady state condition in which input and output rates are equal. The loss processes are degrading reactions and advection. Unlike the Level II model, equilibrium between media is not assumed and, in general, each medium is at a different fugacity. A mass balance applies not only to the system as a whole, but to each compartment. Rates of intermedia transport are calculated using D values which contain information on mass transfer coefficients, areas, deposition and resuspension rates, diffusion rates, and soil runoff rates. It is now essential to define inputs to each medium separately, whereas in Level II only the total input rate was requested.

Mass balances are calculated for the four bulk media of air (gas + aerosol), water (solution + suspended sediment + biota), soil, (solids + air + water), and sediment (solids + pore water). Equilibrium exists within, but not between media. For example, sediment solids and pore water are at equilbrium, but sediment is not necessarily at equilibrium with the overlying water.

Physical-chemical properties are used to quantify a chemical's behaviour in an evaluative environment. Three types of chemicals are treated in this model: chemicals that partition into all media (Type 1), involatile chemicals (Type 2), and chemicals with zero, or near-zero, solubility (Type 3). The model can not treat ionizing or speciating substances. The Level III model assumes a simple, evaluative environment with user-defined volumes and densities for the following homogeneous environmental media (or compartments): air, water, soil, sediment, suspended sediment, fish and aerosols.

This model gives a more realistic description of a chemical's fate including the important degradation and advection losses and the intermedia transport processes. The distribution of the chemical between media depends on how the chemical enters the system, e.g. to air, to water, or to both. This mode of entry also affects persistence or residence time.

Three persistences are calculated, an overall value, T_O, and individual persistences attributable to reaction only, T_R, and advection only, T_A. Note that $1/T_O$ equals the sum of $1/T_R$ and $1/T_A$.

The rates of intermedia transport are controlled by a series of 12 transport velocities. Reaction half-lives are requested for all 7 media. The advective residence time selected for air also applies to aerosols and the residence time for water applies to suspended sediment and fish. The advective residence time of aerosols, suspended sediment and fish cannot be specified independently of the air and water residence times.

Features of the Level III Program:
Provides a database of chemicals and chemical properties.
Permits temporary additions/changes of chemicals and their properties to a simulation.
Permits permanent additions, changes and deletions of chemicals and their properties to the database.
Provides a database of environments and environmental properties.
Permits temporary additions/changes of environments and their properties to a simulation.

Permits permanent additions, changes and deletions of environments and their properties to the database.
Provides context-sensitive Help.
Displays and prints the Level III model calculations, as performed by the program.
Allows the printing of simulation tables, the summary diagram, and a small selection of charts.
Allows the program results to be saved as a comma separated value file, readable by spreadsheet programs such as Excel.

This program was based on the following publication:

Mackay, D.2001. "Multimedia Environmental Models: The Fugacity Approach - Second Edition", Lewis Publishers, Boca Raton, pp.1-261.

The required input data are:
Chemical Properties:

- chemical name
- molecular mass
- data temperature
- reaction half-life estimates for

- air
- water
- soil
- sediment
- aerosols
- suspended sediment
- aquatic biota

- Type 1 chemicals

- water solubility
- vapour pressure
- log Kow
- melting point

- Type 2 and 3 chemicals

- partition coefficients

Environmental Properties:

- areas and depths for all bulk media
- volume fractions for all subcompartments
- densities for all subcompartments
- organic carbon content (soil, sediment & suspended sediment only)
- fish lipid content (Type I chemicals only)
- advective flow residence times for air (including aerosols), and water (including suspended sediment and aquatic biota)
- advective flow residence time for sediment burial
- transport velocities

- air side air-water mass transfer coefficient - water side air-water mass transfer coefficient - rain rate - aerosol deposition velocity (wet and dry combined) - soil air phase diffusion mass transfer coefficient - soil water phase diffusion mass transfer coefficient - soil air boundary layer mass transfer coefficient - sediment-water mass transfer coefficient - sediment deposition velocity - sediment resuspension velocity - soil water runoff rate - soil solids runoff rate

Emissions:

- chemical input rates for each bulk medium or compartment
- inflow concentrations in air and water

Model Output:

- partition coefficients (Type 1)
- Z values
- fugacity of each medium
- intermedia transport rates and D values
- reaction and advection D values and loss rates
- residence times or persistences (overall, reaction, and advection)
- concentrations and amounts for each medium
- a summary diagram
- charts of key results

This program is only available in compiled form. A "readme.txt" file with more detailed technical information is included in the zipped file.

Minimum sytem requirements:
Pentium-75MHz with 8 Mb of RAM running Windows 95. On some systems it may be necessary to adjust your screen resolution.

The Level III Model Version 2.70, released March 2002 continues to be available.

For non-Windows users the BASIC, evaluative, Level I, II and III fugacity models are available.

Please read the **LEVEL III SOFTWARE LICENSE** before downloading the software. Use of the software constitutes your agreement to abide by the terms and conditions set out in the license agreement.

Bug Report

Register to Download

CEMC Home | General Information | Models | Graduate Studies | Contact

Last updated July 26, 2004.

Canadian Environmental Modelling Centre

Level III Software License

Version 2.80 - May 2004

The Level III software program is provided to interested parties at no cost. We do request that you provide us with registration information at the time of download from our website for our own information and statistical purposes.

The use of the Level III software program is governed by the following legal agreement. The purpose of the agreement is to ensure that all users are treated equitably, and we are not disadvantaged in any way by providing the software. All use of this software is conditional upon your compliance with the license terms which follow. If you do not agree to the terms of this license agreement, or do not comply with the terms and conditions of this agreement, you are not permitted to use this software and are required to remove the Level III software program from your computer system, and destroy all copies of the software. The use of the Level III software program constitutes your agreement to abide by the terms and conditions set out in this document.

THE Level III SOFTWARE PROGRAM, HEREIN CALLED THE "SOFTWARE", IS OWNED BY TRENT UNIVERSITY AND IS PROTECTED BY CANADIAN COPYRIGHT LAWS. UPON YOUR AGREEMENT TO AND COMPLIANCE WITH THE TERMS OF THIS LICENSE AGREEMENT, TRENT UNIVERSITY GRANTS YOU, HEREIN CALLED THE "LICENSEE", THE FOLLOWING NON-TRANSFERRABLE, NON-EXCLUSIVE RIGHTS OF USE. TRENT UNIVERSITY HAS THE RIGHT TO TERMINATE THIS AGREEMENT IF THE "LICENSEE" FAILS TO COMPLY WITH ANY TERM OR CONDITION OF THIS AGREEMENT. NO TITLE TO THE INTELLECTUAL PROPERTY IN THE "SOFTWARE" IS TRANSFERRED TO YOU. THE "LICENSEE" DOES NOT ACQUIRE ANY RIGHTS TO THE "SOFTWARE" EXCEPT AS EXPRESSLY SET FORTH IN THIS LICENSE.

GRANT OF LICENSE REGARDING THE Level III PROGRAM

TRENT UNIVERSITY GRANTS THE "LICENSEE" THE FOLLOWING RIGHTS REGARDING THE USE OF THE "SOFTWARE":

1) USE OF THE "SOFTWARE" FOR THE "LICENSEE'S" PERSONAL OR BUSINESS PURPOSES.

2) COPYING THE "SOFTWARE"
i) THE "LICENSEE" MAY NOT MAKE COPIES OF THE "SOFTWARE" OTHER THAN THOSE GRANTED BY LAW FOR ARCHIVAL OR BACKUP PURPOSES.
ii) THE "SOFTWARE" MAY BE TRANSFERRED TO THE HARD DISK OF ANY COMPUTER, OR NETWORK OF COMPUTERS, BELONGING TO THE "LICENSEE".

RESTRICTIONS REGARDING THE Level III PROGRAM

1) THE "LICENSEE" MAY NOT REVERSE ENGINEER, DECOMPILE, DISASSEMBLE, MODIFY, TRANSLATE, OR ALTER THE SOFTWARE AND/OR THE ASSOCIATED FILES, OR IN ANY MANNER SUPPORT OR CAUSE SUCH TO OCCUR.

2) THE "LICENSEE" MAY NOT DISTRIBUTE, SUBLICENSE, LEASE, SELL, RENT OR OTHERWISE TRANSFER THE "SOFTWARE", OR ANY MODIFICATION OR DERIVATIVE THEREOF, TO ANY OTHER INDIVIDUAL OR GROUP FOR ANY REASON. (Additional copies can be obtained directly from this web site.)

LIMITATIONS OF LIABILITY AND DISCLAIMER OF WARRANTY

THERE ARE NO WARRANTY RIGHTS GRANTED TO YOU, THE "LICENSEE", REGARDING THE "SOFTWARE". THE "SOFTWARE" AND ACCOMPANYING WRITTEN MATERIALS ARE SUPPLIED TO THE "LICENSEE" "AS IS" WITHOUT WARRANTY OF ANY KIND. TRENT UNIVERSITY DOES NOT GUARANTEE, WARRANT, OR MAKE ANY REPRESENTATIONS, EITHER EXPRESSED OR IMPLIED, REGARDING THE USE, OR THE RESULTS OF THE USE OF THE THE "SOFTWARE" OR THE Level III WRITTEN MATERIALS WITH REGARDS TO RELIABILITY, CURRENTNESS, ACCURACY, CORRECTNESS, OR OTHERWISE. THE "LICENSEE" ASSUMES THE ENTIRE RISK AS TO THE RESULTS AND PERFORMANCE OF THE "SOFTWARE".

TRENT UNIVERSITY SHALL NOT BE LIABLE UNDER ANY CIRCUMSTANCES, FOR ANY DAMAGES WHATSOEVER, ARISING OUT OF THE USE, OR THE INABILITY TO USE, THE "SOFTWARE", EVEN IF TRENT UNIVERSITY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

Model Description

CEMC Home | General Information | Models | Graduate Studies | Contact

Last updated May 25, 2004.

Canadian Environmental Modelling Centre

Level III Software Update

Version 2.80 - May 2004

Since the Version 2.70 (March, 2002) release:

Simulation ID

- There is now an input box for additional notes which is then entered directly into the print simulation section and will appear on the final printout.

Chemical Properties

- There was an error in v. 2.70 which did not allow changes in the Chemical Properties and New Chemicals sections to be saved to the Database. This has been corrected and applies to all three types of Chemical.
- There is a change to the half-life checkboxes. When the half-life checkbox is ticked so that the half-life is considered to be negligible and then the user enters a new half-life into the input box the check box is unchecked and the new value is stored.
- In all input boxes on the Chemical Properties and New Chemical pages, when errors are caught the field in error is highlighted and cleared in order for the user to reenter an appropriate value.
- The Henry's Law Constant display has been removed from this form and is now only viewable in results. This was done to ensure a smoother run of the program. It will not affect the user's results in any way.

Environmental Properties

There is an increased flexibility for defining Environments.

- A water area of 0 m2 is now allowed. When this value is entered it will automatically change the water and sediment depths to 0 m.
- The areas of soil and sediment are no longer displayed on this input form.
- It is now impossible to enter a zero value for area or depth of air.
- There are more Standard Environments to choose from. The single Default Environment has been replaced by Default 1, 2 and 3. Three additional environments are included as described below. The user is not permitted to change the properties of these environments in the database. - Default 1 is similar to the Default Environment in v. 2.70 except that the air residence time has been changed to reflect a more realistic value. - Default 2 is based on the Southern Ontario region defined in ChemCAN v 6.00. - Default 3 is similar to the Shield Lake Region. - Shield Lake Region (Mackay, D., Webster, E., Woodfine, D., Cahill, T., Doyle, P., Couillard, Y., Gutzman, D. 2003. Towards Consistent Evaluation of the Persistence of Organic, Inorganic and Metallic Substances. Human and Ecological Risk Assessment (HERA). 9: 1445-1474.) - EQC - standard environment (Mackay, D., Di Guardo, A., Paterson, S., Cowan, C.E. 1996. Evaluating the Environmental Fate of a Variety of Types of Chemicals Using the EQC Model. Environ. Toxicol. Chem. 15: 1627-1637.) - Beyer Environment (Beyer, A., Mackay, D., Matthies, M., Wania, F., Webster, E. 2000. Assessing Long-range Transport Potential of Persistent Organic Pollutants. Environ. Sci. Tech. 34: 699-703.) Additional entries in the database can

be changed and are thus are the responsibility of the user.
- There was an error in v. 2.70 which did not allow changes in the Environmental Properties and New Environments sections to be saved to the Database. This has been corrected.
- Labeling on the Environmental Properties' tabs is now consistent, where titles and units are given only where necessary. There has been to reduce redundancy and to ensure that all appropriate units are displayed.
- It is now possible to enter a value of zero into any of the volume fraction boxes as long as the total value of the boxes is equal to one or another appropriate value.
- There is now a Help Button on the New Environment Form.

Emissions

- There are no changes to this section.

Chemical Parameters

- There are no changes to this section.

Environmental Parameters

- There are no changes to this section.

Results

- The Advection tab has been corrected and now includes the air residence time as defined in Chemical Properties by the user.
- The Results section has been corrected to correctly display the D-value units, and aquivalence units for Type 2 chemicals.
- A more balanced and easy to follow look has been given to the Results section.

Diagram

- When the diagram is closed the user's cursor now advances naturally onto the Charts button.

Charts

- Bar charts are used for the fugacities, masses and concentrations in each medium. Pie charts are used for relative amounts in each medium.

Error Checking

- The error message for a user entering a value that is Outside Reasonable Bounds has been updated so that it now states "The absolute value of the exponent is too large. Please enter a reasonable value".
- On the Densities Tab in Environmental Properties the appropriate error message is now displayed when a zero value is entered.
- Error checking has been corrected so that values larger than and smaller than the absolute value of 3e?38 are caught and the program does not crash.
- Error checking has been corrected so that when the area of water > area of air the appropriate messages are displayed and corrections can be input.

Calculations

- Z-value calculations for Type 3 chemicals has been corrected so that it displays the correct values for pore water in sediment.

Print

- The option for printing the date and time when printing Tables has been corrected so that it works.
- Note: All corrections to units and calculations in above sections have carried over to the printout.

Help Files

- The Environmental Properties Help File now contains a discussion of the database including source information.
- The Environmental and Chemical Database Operations' Help Files are now applicable to both types of operations.
- The Results Help File now contains an updated definition of fugacity, consistent with Level III, steady-state non-equilibrium, concepts.

About

- The About Button on the Main form has been altered so that it no longer includes the program's name.
- The New Features Button, which you have already used, is a New Feature to this program.

Since the Version 2.65 (February 2002) release:

- The chemical half-lives are now correctly saved to the database.
- It is now possible to add new Type 2 and 3 chemicals using the "New Chemical" button.
- The name of the environment is now displayed under the chemical name before printing the results.
- Added Chart display of selected results. These can also be printed with the rest of the results.
- Corrected the description of the save-to-file function in the general Help file.
- Updated the Help file for the Print function.

Since the Version 2.20 (December 1999) release:

- The aerosol deposition velocity parameter is now entered as the dry deposition velocity and a scavenging ratio rather than as a combined wet and dry velocity.
- Updated terminology to be consistent with recent work.

Suspended sediment is now more correctly referred to as suspended particles and "pure air" is now referred to as air vapour.
Molecular mass is now more correctly referred to as molar mass.

- An over-sight in the numerical formatting routine was corrected.
- Corrected model name in various Help files and in the readme.txt file.
- Calculations now print correctly.
- All model output, whether viewed on screen, printed, or saved to a file, are identical except that values in the file are not formatted. This allows the user to see the whole number calculated by the model without any rounding effects.
- Removed irrelevent D value totals.
- Improved error checking.
- Incorrect information in the Help files for the Environmental Properties and Emissions input forms

has been removed.

- In the display of results, the layout of the chemical parameters was improved.
- The aerosol-air partition coefficient is no longer displayed for Type 2 chemicals where it is not applicable.
- A label error in display of Intermedia Transport was corrected. "Water to soil" is now "Soil to air".
- Concentration (ug/g) in sediment solids was wrongly displayed as the value for soil pore air.
- Overall D values have been added to Results display.

- Results are now saved to a comma separated value file making it more generally accessible to a variety of spreadsheet programs.
- The units of Fugacity for Type 1 and 3 chemicals were corrected in the save-to-file output.
- The save-to-file output was recording the sediment D value for reaction as that of soil. This has been corrected.
- A listing of environment properties used were added to the save-to-file output.
- A date and time stamp was added to save-to-file.
- Layout of the save-to-file output was generally improved.

- All input values are italicized in the printed output to facilitate future simulation duplication.
- The layout of the printed output of chemical parameters was improved.
- The aerosol-air partition coefficient is no longer printed for Type 2 chemicals where it is not applicable.
- The aerosol-water partition coefficient is no longer printed for Type 1 where it is unnecessary and illogical.
- The "Save Diagram" option was removed. It did not work consistently. The diagram can be captured using the "Print Screen" button on the keyboard, and pasted into a file to be saved. A note was added to the Help file for the Diagram to indicate this.

- A date and time stamping option was added to the printed output.

Special thanks to Jenn Brimacombe, Angela McLeod, Chris Warren and the WE512 class of 2001 for checking the model results.

Since the Version 2.10 (September 1999) release:

- The values displayed in the Diagram for concentration in soil and sediment have been corrected. Previously there was an error in the unit conversion.
- The units listed for Fugacity in the Tables of Printed Results have been corrected.
- The typographical error in the "note" about the air and water densities was corrected.

Model Description

CEMC Home | General Information | Models | Graduate Studies | Contact

Last updated May 25, 2004.

E_TX 1.3a

A Program to Calculate Confidence Limits for Hazardous Concentrations Based on Small Samples of Toxicity Data

Tom Aldenberg

rivm
National Institute of Public Health &
Environmental Protection
Bilthoven, The Netherlands

May 1993 719102015

Contents

I. Introduction

Welcome to ***E_TX***, the Ecotoxicological Extrapolation Program from **rivm**!
E_TX is a computer program running on MS-DOS computers. *E_TX*, short for EcoToX, currently version *1.3a*, handles the extrapolation of laboratory toxicity data to values, that may be of interest to policy makers in setting standards for environmental protection. The program may also be of use in other areas, e.g. in human health-oriented problems.
As a short motivating example, suppose one is confronted with the next seven NOEC (No Effect) concentrations for Cadmium for various different soil fauna species:

```
0.97   3.33   3.63   13.5   13.8   18.7   154
```

in some unit. These are real data from Van Straalen & Denneman (1989), who adapted this extrapolation technique from Kooijman (1987). Then, *E_TX* calculates **`0.53`** as the estimate of the 5th percentile of the hypothetical statistical distribution from which the data are thought to derive. This 5th percentile is the so-called *hazardous* concentration, above which 95% of the species seems relatively safe. We also speak about 95% species protection. This hazardous concentration is also indicated as the HC_5.
The estimate of the hazardous concentration just employed is a so-called *median* estimate: if everyone would calculate this estimate for similar batches of seven data in the same manner, for instance by using *E_TX*, then the median of the distribution of answers, of which **`0.53`** is one particular instance, would equal the hazardous concentration.
E_TX also calculates a second value, **`0.03`**, for this example, that, if everyone

would do so for their data, would result in a distribution of answers of which the 95th percentile would equal the hazardous concentration. That is, we are *confident* that we *under*estimate the hazardous concentration by 95%. One can either think of this estimate as the *left confidence limit* of a 90% double-sided confidence interval, or as the one-sided 95% confidence underestimate.

The basic idea, henceforth, is that laboratory species display different sensitivities with regard to the adverse effects of a particular toxic substance, as expressed by NOEC concentrations or LC_{50} concentrations. If nothing is assumed mechanistically, these species NOECs, or whatever, are thought to derive from some statistical distribution. 'Extrapolation', as it is called, amounts to estimating percentiles of this distribution with a certain confidence from a perhaps small set of toxicity data.

The statistical theory behind extrapolation to percentiles is treated in Erickson & Stephan (1985), Kooijman (1987), Van Straalen & Denneman (1989), Wagner & Lokke (1991), and Aldenberg & Slob (1993). *E_TX* is conceived as a tool for the statistical analysis of toxicity data sets. Although *E_TX* is relatively user-friendly, and can be run by decision makers, it is not specifically designed as a decision support system for setting environmental standards. *E_TX* does not care about what the nature of the data is, one feeds to it. It may be used on NOEC concentrations, EC_5 concentrations, LC_{50} concentrations, etc. In fact, the data may not be toxicity data at all, but one should be aware of the fact that the data are always log transformed. The toxicity data may pertain to any taxon level: species, genera, or even higher taxa (e.g. Crustaceans, Mammals, etc.). This means that a toxicity data set may consist of a batch of species toxicity data, or a batch of toxicity data, one for each genus, a batch of phylum data, and so on.

Nor does *E_TX* know about any specific environmental protection terminology, such as permissible risk levels, maximum allowable concentrations, safe or reference or background concentrations. *E_TX* is a program to experiment with different toxicological data sets, different confidence levels, and different species protection levels. Right now, it handles three types of statistical distribution, i.e. the log-logistic, the log-normal, and the log-triangular, one species protection level (95%), and two levels of confidence of underestimation (95% and 50%). But these may be extended in future versions. Hence, *E_TX* may develop over time, as more analysis tools are incorporated. We would welcome any user remarks, that could lead to improvement.

The reader who wants more information about the practical and theoretical considerations in a decision makers framework is referred to Slooff (1992) and OECD (1992).

Next to the estimation of hazardous concentrations from laboratory toxicity data,

or *extrapolation*, initial steps have been taken in E_TX to incorporate the estimation of species hazard at given environmental or experimental concentrations, here called *hazard assessment*. Whereas extrapolation goes from a pre-set species protection level and a batch of toxicity data to concentrations to be declared as environmental standard, or objective, hazard assessment goes from current or predicted environmental concentrations to estimated species protection levels. Here also, a confidence approach would be implied, but more work has to be done on that. Extrapolation is treated in the literature cited, but hazard assessment, as defined here in a statistical extrapolation-oriented framework, is treated in Appendix A of this Manual.

Acknowledgments.

I would like to thank *Kees van Leeuwen* for being the motor behind this program, and for his incredible patience; *Hans Canton, Wilbert Slooff,* and *Wout Slob* for their stimulating interest. *Esther Guinée, Janneke Hoekstra, John Janus, Robert Luttik, Erik van de Plassche, Marieke Polder,* and *Theo Traas* commented on earlier versions of the program. The correspondence with *Nico van Straalen* is acknowledged. *Maarten 't Hart* helped with the final testing and documentation.

TA

II. User's Guide

A. Installation

E_TX can be run on MS-DOS Personal Computers, or so-called Compatibles, right out of the box from floppy disk. So, strictly speaking, there is no obligation to install E_TX on a hard disk. Obviously, it is more convenient to do so.
Installation under MS-DOS is very simple. Be sure, you see the DOS prompt **`C:>`** after starting the PC. If wanted, make a directory to place E_TX in, by typing:

```
MD ETX
```

or some other directory name. Go to this directory:

```
CD ETX
```

Put the E_TX floppy disk in drive **`A:`** or **`B:`**. Copy all the files from the E_TX floppy disk to the hard disk (**`C:`**) by typing:

```
COPY A:*.*
```

or:

```
COPY B:*.*
```

After the files have been copied, remove the floppy disk from the floppy disk unit.
Now you are ready to run *E_TX* from the hard disk.

B. Running *E_TX*

Running *E_TX* is even simpler. If the file **ETX.EXE** is in the current directory, just type:

```
ETX
```

Otherwise, first go to the directory, where **ETX.EXE** is located. If the MS-DOS PATH is set by you, or someone else, to scan the **ETX-directory** for commands, then typing **ETX** will work from any directory. In the current version, data files that are saved by, or are to be used by *E_TX* are put into the same directory, as where the program runs. This might change in a future release.
E_TX is operated through a menu system. You may be familiar with menu systems by other programs. If not, learning to operate the *E_TX* menu system should not cause any problem. 'Menu' options constitute alternative choices for the user of a program to control a program.
Menu items are displayed vertically in *E_TX*. There may be ten items, but usually less. Each list of menu options makes a menu screen. Menu options ending with a slash (**/**) lead, when activated, to a submenu containing new options. Those ending with a period (**.**), when activated, cause some action to be taken by *E_TX*. Options within parentheses, if any, are currently not implemented.
The main menu screen of *E_TX* after starting the program reads:

```
ETX 1.3a:

     1. Data/
     2. Statistics/
     3. Extrapolation/
     4. Hazard/
     5. Results/
     6. Quit/

[Edit Keys] to Select; [Enter] to Activate:
```

All options have trailing slashes, so each choice, when activated, gives rise to a new menu. These six options form a logical sequence of doing an extrapolation analysis: getting data, study them, carry out extrapolating exercises

and hazard assessment, collect output, and quit the program. But they need not be chosen in this exact order. If you try to do illogical things, however, such as extrapolation without entering toxicity data, E_TX will complain.
One can activate a menu item in either of two ways. The first is by pressing the number that precedes the item. E_TX should immediately respond to that. Or, secondly, one can use the arrow keys, or edit keys: **[Down]**, **[Up]**, and so on, to highlight next and previous options. In this case, no action is taken, until you press the **[Enter]** key. The last line of the menu screen, the Active Key bar, indicates what keys are active. On the screen and in this manual, brackets denote a key to be pressed, not a sequence of characters. So, **[Enter]** is the Enter or Return key.
When you are in a submenu, or sub-submenu, pressing **[Esc]**, or **[PgUp]**, brings you up one level. Pressing **[F10]** brings you up to the main menu from anywhere in the tree. Pressing initial characters does *not* activate a menu item.
Above the menu options, you find the name of the program (**ETX**) and its current version number (**1.3a**). If it says also something like **Beta**, followed by a number, then you do not have an official release, but a version distributed for review.
This title bar of a menu screen is also used to indicate where the user is located in the menu hierarchy. When you are not at the main, or top menu level, then previous menu choices have been added to the program and version indicator, separated by slashes (**/**). This results in a representation of the path followed through the menu hierarchy (tree), similar to the way operating systems display directory hierarchies. In this way, going down the menu tree (if main menu is thought to be the highest level) makes the menu path string to grow, while going back up the hierarchy does make it shrink, until you are back at the top level, and only the program and version indicator remain.
Hence, if the path string at the top of a menu screen reads:

```
ETX 1.3 a/ Statistics/ Logarithms/ Toxicity:
```

one has apparently chosen option **2. Statistics** from the E_TX main menu, then **1. Logarithms** from the Statistics menu, then **1. Toxicity** from the Logarithms menu, and one is currently facing the two options and Active Key bar:

```
    1. As Is.
    2. Sorted.

[Edit Keys] to Select; [Enter] to Activate; [Esc] to Quit:
```

These options have trailing periods, so they will do something for you. Option

1 will show you a list of the $\log_{10}$s of the toxicity data, if there are any, in the order as they have been entered. Option 2 will do the same thing, only in sorted order from smallest to largest. Since the action follows immediately, and the menu screen will be cleared, these action options are not added to the menu path string anymore. Note the extended Active Key bar, with **[Esc] to Quit** added, indicating that **[Esc]** (the Escape key) brings you back up.
In the Reference Manual, we will represent subsequent menu choices by the path traveled through the menu tree, as given by the menu title bar, but with two slight modifications: the leading program version indicator, **ETX 1.3a**, is abbreviated to **ETX**, and last menu options with trailing periods that lead to an action are added to the path. So, referring to the previous example, the act of listing the toxicity data logarithmically in the original order is effectuated through this sequence of menu choices:

ETX/ Statistics/ Logarithms/ Toxicity/ As Is.

From now on, we refer to such strings as the menu path, or path, of a command. Uncomplete paths, ending with a slash, are also called paths.

C. Leaving E_TX

As soon as you are done, you can quit (leave) E_TX from main menu through the menu choices: **ETX/ Quit/ Leave ETX!** Note that we have used the path indication method explained in the previous paragraph. On the screen, menu option **Leave ETX** has a trailing exclamation mark, which indicates that there is no mercy in case you have not saved the data that you may have entered. There is *no extra warning* if you haven't done so!
You can find out about the status of the data under **ETX/ Data/ Show/**, that will display the current data set, as well as inform the user, where the data came from, and whether they were saved.
A very crude way to leave the program at any point in the menu hierarchy is pressing **[Ctrl]+[End]** and then **[Ctrl]+[Y]**. If you miss **[Ctrl]+[Y]**, you are back where you were. Use this only in a great hurry.
A real crash route is **[Ctrl]+[Break]**, which aborts the program at once.

III. Reference Manual

The Reference Manual consists of two parts:

A. an overview of the *E_TX* Menu hierarchy or tree, including a short description of what each option means or does, and
B. an alphabetical Reference list for lookup of terms relating to *E_TX*: menu options and general issues like Files, Version, and so on.

A. Menu tree

The *E_TX* menu system is a tree-like hierarchy. At the so-called root of the tree, usually considered the highest level of the menu, we have *E_TX* itself. This level is activated by typing **ETX** at the MS-DOS prompt. After clearing the opening screen, we arrive at the first level down the tree, with six branches. The menu screen reads:

```
ETX 1.3a:

1. Data/
2. Statistics/
3. Extrapolation/
4. Hazard/
5. Results/
6. Quit/
```

followed by the Active Key bar. The **Data** menu leads, when activated (see User's Guide), to another menu screen at depth two. This reads:

```
ETX 1.3a/Data:

1. Enter/
2. Read/
3. Show/
4. Edit/
5. Save/
```

Note the growing menu path at the first line. Now, activating menu option **Enter** leads to just another menu screen at depth three.
Below follows a complete representation of the tree with both the layout and decimal numbering indicating the hierarchical structure. Each option keyword is shortly explained between parentheses. If an option does not have a submenu, then it denotes some action to be taken by *E_TX*, often leading to further prompts for user action, like entering or modifying data, actions to save data in disk files, listing of results, statistical output, etc.

E_TX menu tree:

```
1 Data/              (data management)
     1.1 Enter/           (enter new data through the keyboard)
          1.1.1 Toxicity.      (enter laboratory toxicity data)
          1.1.2 Exposure.      (enter environmental exposure data)
     1.2 Read/            (read data from files)
          1.2.1 Toxicity.      (read toxicity data)
          1.2.2 Exposure.      (read exposure data)
     1.3 Show/            (look through the data)
          1.3.1 Toxicity/      (show toxicity data)
               1.3.1.1 As Is.       (in the original order)
               1.3.1.2 Sorted.      (in sorted order)
          1.3.2 Exposure/      (show exposure data)
               1.3.2.1 As Is.       (in the original order)
               1.3.2.2 Sorted.      (in sorted order)
     1.4 Edit/            (modify entered or read data)
          1.4.1 Toxicity.      (edit toxicity data)
     1.5 Save/            (save data to files on disk)
          1.5.1 Toxicity/      (save toxicity data)
               1.5.1.1 As Is.       (in the original order)
               1.5.1.2 Sorted.      (in sorted order)
          1.5.2 Exposure/      (save exposure data)
               1.5.2.1 As Is.       (in the original order)
               1.5.2.2 Sorted.      (in sorted order)
2 Statistics/        (look at statistical data evaluations)
     2.1 Logarithms/      (show log10 transformed data)
```

```
          2.1.1 Toxicity/     (logarithmic toxicity data)
               2.1.1.1 As Is.      (in the original order)
               2.1.1.2 Sorted.     (in sorted order)
          2.1.2 Exposure/     (logarithmic exposure data)
               2.1.2.1 As Is.      (in the original order)
               2.1.2.2 Sorted.     (in sorted order)
     2.2 Basic Statistics/(look at simple data summaries)
          2.2.1 Toxicity/     (of the toxicity data)
               2.2.1.1 Logistic.   (the logistic distribution)
               2.2.1.2 Normal.     (the normal distribution)
     2.3 Goodness-of-Fit/(Goodness-of-Fit of distributions)
          2.3.1 Toxicity/     (fitted on the toxicity data)
               2.3.1.1 Logistic.   (the logistic distribution)
               2.3.1.2 Normal.     (the normal distribution)
3 Extrapolation/    (look at extrapolation results)
     3.1 Logistic.       (for the logistic distribution)
     3.2 Normal.         (for the normal distribution)
     3.3 Triangular.     (for the triangular distribution)
4 Hazard/           (look at hazard assessment results)
     4.1 Logistic/       (for the logistic distribution)
          4.1.1 Fixed Hazard.(exposures at fixed hazard levels)
          4.1.2 At Exposure/  (at the exposure data)
               4.1.2.1 As Is.      (in the original order)
               4.1.2.2 Sorted.     (in sorted order)
5 Results/          (collect the results)
     5.1 Save/           (save results on disk)
          5.1.1 All.          (save all results)
6 Quit/             (quit ETX)
     6.1 Main Menu.      (no, back to main)
     6.2 Leave ETX!      (yes, definitely leave ETX)
```

B. Reference List

The Reference List serves as an explanatory alphabetic lookup table for E_TX related keywords. Among these are all menu options, and features of more general interest to the E_TX user. The Reference List is both indexed in the Table of Contents, as well as in the General Index at the back of the Manual. Some of the entries refer to related issues elsewhere in the list. If a keyword is a menu option, then the relevant menu paths are given first. The meaning and notation of menu paths, or paths for short, is explained in the User's Guide (Running E_TX).

1. All

Path: **ETX/ Results/ Save/ All.**

All is the only option under **ETX/ Results/ Save**, indicating that one can only save all the results to a file, not a selection of them. **All** prompts for a file name (default extension **.ETX**). One can edit the line that contains **.ETX** pre-written, including the extension. If no extension is given, extension **.ETX** is added by *E_TX*. Existing files can never be overwritten by *E_TX*.

2. As Is

Paths:
ETX/ Data/ Show/ Toxicity/ As Is.
ETX/ Data/ Show/ Exposure/ As Is.
ETX/ Data/ Save/ Toxicity/ As Is.
ETX/ Data/ Save/ Exposure/ As Is.
ETX/ Statistics/ Logarithms/ Toxicity/ As Is.
ETX/ Statistics/ Logarithms/ Exposure/ As Is.
ETX/ Hazard/ Logistic/ At Exposure/ As Is.

As Is always refers to the original order in which the data have been entered or read from a file. As can be seen from the above paths, **As Is** refers to the last menu screen to fire the complete command. On this last menu screen, it is always the first option of two, the second being **Sorted**. Hence, pressing **[Enter]** activates it as the default. What **As Is** actually does, depends on the previous menu options, to which the reader is referred..

3. At Exposure

Path: **ETX/ Hazard/ Logistic/ At Exposure/**

At Exposure is the second option of **ETX/ Hazard/ Logistic/**. It displays the estimated hazard at the entered or read exposure concentrations, given the logistic density estimate fitting the current toxicity data set. See: **Hazard**. The alternative option of **At Exposure** is **Fixed Hazard** (see there, and see **Hazard**, as well as Appendix A for the full story).
The two options of **At Exposure** are: **As Is** and **Sorted**. These refer to the on-screen representation order of the data, i.e. unsorted or sorted.

4. Basic Statistics

Path: **ETX/ Statistics/ Basic Statistics/**

Basic Statistics is the second option of **Statistics**. Most of the statistics shown on screen are used in the extrapolation and hazard calculations. There is no need to study or record these statistics, unless you want to check the extrapolation and hazard estimates by hand, study the influence of outliers on means and standard deviations, see where the data are located in log space, and so on.

Basic Statistics has only one option: **Toxicity**. No statistics are calculated in E_TX for the exposure data set. The exposure data are just used to estimate the hazard at each individual exposure value. The statistics are displayed with respect to two distributions: the logistic and the normal distribution. Both screens show the toxicity sample statistics: mean and standard deviation, and the respective extrapolation constants.

These extrapolation constants are given in Tables B1 (logistic) and B2 (normal). Table B1 is derived from Aldenberg & Slob (1993). Table B2 (95% confidence) is derived from Wagner & Lokke (1991). The median extrapolation constants for the normal distribution (Table B2: median) have been kindly provided by dr. R.J. van Wijk (AKZO Research CRL, Arnhem). These constants have been determined through simulation by drawing 5000 random samples for each sample size. We have not been able to check the performance of these constants, but they seem to match the median extrapolation constants for the logistic distribution quite well. No such table seems available in the statistical literature. Extrapolation constants for sample sizes not in Table B1 or B2, are estimated in E_TX by linear interpolation.

The logistic basic statistics screen presents some parameter estimates of the logistic distribution parameters, α and β. It is important to note that all are $\log_{10}$ values. Hence, they should be compared with the values as given under **ETX/**

Statistics/ Logarithms/. The moment estimates are those of Kooijman (1987):

$$\hat{\alpha} = \bar{x}$$

and

$$\hat{\beta} = s_n \cdot \frac{\sqrt{3}}{\pi}$$

In fact, for a toxicity data set of size n, s_n is the bias-corrected sample standard deviation on the basis of (n - 1). Hence, one could speak of bias-corrected moment estimates.
The maximum likelihood estimates of α and β are more difficult to estimate. We need them for the goodness-of-fit calculations for the Logistic distribution. Now, α and β have to be solved from the nonlinear equations:

$$\sum_i \frac{1}{1 + \exp((x_i - \alpha)/\beta)} - \frac{n}{2} = 0$$

and

$$\sum_i \frac{(1 - \exp((x_i - \alpha)/\beta)).(x_i - \alpha)/\beta}{1 + \exp((x_i - \alpha)/\beta)} + n = 0$$

The equations are given by D'Agostino & Stephens (1985). We did check out, that they indeed follow from a maximum likelihood argument. These equations must be solved numerically. We have used a discrete Newton-Raphson procedure to do so. A numerical subtlety turned out to be those cases where the x_i are symmetrically located around their average, for example: -1 and +1. Then the first equation with α equal to the x-average is uniformly satisfied for all β, which blows up the iterative procedure. We decided to put α equal to x-average, in these cases, and to iterate only the second equation to solve for β.
The logistic basic statistics screen further displays so-called HC_5 fitting estimates for the logistic parameters. These are described in Appendix A. See also **Hazard**. The idea is that, if an exposure value happens to be equal to the median (extrapolation) estimate of the HC_5, on the basis of the toxicity data,

then the estimated hazard at that exposure concentration should be 5%.Appendix A explains that this leads to the estimate:

$$\beta = k_L \cdot s_n / C_5$$

The **Basic Statistics** display screen for the Normal distribution contains analogous parameter estimates for location and scale parameter μ and σ. The moment estimates, corrected for bias, are identical to the sample mean and sample standard deviation. The maximum likelihood estimate of σ is the raw standard deviation of the sample, not corrected for bias, i.e. on the basis of *n*, the number of toxicity tests. These are not used by *E_TX*, but only given for completeness. For the normal case, goodness-of-fit calculations are based on the sample mean and sample standard deviations, not on the maximum likelihood estimates, as is the case with the logistic distribution.

5. Batch mode

The previous version of *E_TX*: ETX 1.2A did have a batch mode. You could type under MS-DOS a command line like: **ETX MYDATA.DAT 3**, in order to process the third line of the data in file **MYDATA.DAT**. Then, the menu system was by-passed, and an output file **MYDATA.3** was automatically generated. Hence, data files could consist of different sets of toxicity data, e.g. one line per toxic substance. A MS-DOS **.BAT** file, could consist of several such ETX command lines.
In the current version (*1.3a*), reading and saving of files has been much improved. Now, there is only one toxic substance per file, and different commented toxicity data are on separate lines. Again, a batch mode option would be feasible, but it has not been incorporated again. We would like to know whether there is a need for it.

6. Data

Path: **ETX/ Data/**

Data is the first option of the *E_TX* main menu screen. It is analogous to the first entry 'File' on the menu bar of many programs. **Data** gives rise to a submenu. One can enter data interactively through the keyboard (**Enter**), read data from disk files (**Read**), look through data, that were entered or read before (**Show**),

modify data (**Edit**), and save data to disk files (**Save**). The reader is referred to these options for further information.

7. Edit

Path: **ETX/ Data/ Edit/**

Edit is the fourth option of the **ETX/ Data/** menu screen. **Edit** allows the modification of a data set that has been entered through the keyboard, or read from a disk file. Currently **Edit/** has only one option: **Toxicity**. In this version, it is not possible to edit the exposure data from inside *E_TX*.
It is important to keep in mind, that **Edit** is not capable of editing data sets in the way an ASCII editor, or wordprocessor can do that. **Edit** is meant for correcting typing errors in data just entered, or to add comment strings to poorly annotated data; it can be used for sensitivity analysis, e.g. by changing data a little and calculating difference quotients, etc. For instance, it is not possible, to delete data, or to add extra data to the set, or extra comment lines, through **Edit**.
Data modifications, that do change the size of the data set, need an outside full-screen editor, such as the MS-DOS Editor, Turbo Pascal Editor, WordPerfect, or other word processors. In those cases, edit the disk file that contains the data. Hence, for data sets just entered: save them, leave *E_TX*, and start the editor of choice. Save the augmented data to a disk file, then open *E_TX* again and read in the modified file.
For modifications of existing data, **Edit** is fine. Initially, **Edit** works just like **Show** (see: **Show**). One can browse through the data page by page, entry by entry, until the entry to be modified is found, and highlighted, i.e. printed inversely. Then, pressing the **[Space]** bar, opens the possibility to edit the current entry. This Line edit mode is indicated by the extension of the highlight to the full length of the line (full highlight). In this state, the Edit keys (see: Edit keys) are active, and the entry can be modified, or extended. Invalid entries are not accepted. These are non-positive values, values not separated from the comment by white space (spaces and/or tabs), lines without a numeric 'head', and so on. In particular, one cannot enter a comment line starting with an exclamation mark. Pressing the **[Esc]** key, when still in Line edit mode (full highlight), recovers the previous data line.
After completing the modifications, pressing the **[Enter]** key makes the changes permanent. One stays in Browse mode, just as under **Show**, however. So, one can travel through the data again, entrywise, or pagewise, find a new

entry to be modified, press the **[Space]** bar to activate Line edit mode (full highlight), make modifications, press **[Enter]** to confirm, and so on, until you are done.
Now, there are two ways to leave **Edit**: through the **[Esc]** key, or through the **[Enter]** key. The **[Esc]** key restores the old data set, while **[Enter]** confirms all modifications. There are some extra questions asked for confirmation. Hence, there are several occasions in **Edit**, where you can change your mind and restore the previous situation.
One cannot edit exposure data right now. Apart from using an outside editor, there is a work-around in *E_TX*, if you know what you are doing (not recommended though). Read the **.EXP** file as toxicity data through **ETX/ Data/ Read/ Toxicity** (be sure to save the previous toxicity data!). Then, edit, save again, with extension **.EXP**, *E_TX* will not protest against that. The first line of the header of the file written by *E_TX*, however, is erroneous then, since it will read: **! ETX Data File: Toxicity Tests**. However, this information is not seen by *E_TX*, while reading the file as an Exposure data file. It is the responsibility of the user to correct this erroneous comment in the file afterwards with a real-world editor. Afer reading in the modified exposure set, be sure to erase the internal toxicity data set, that is still the exposure set just saved, by reading in a fresh genuine toxicity data set, or by entering new toxicity data.

8. Enter

Path: **ETX/ Data/ Enter/**

Enter is the first option of the **ETX/ Data/** menu screen. **Enter** allows the interactive entry of data through the keyboard, or of new data, overwriting existing data, previously entered, or previously read from a disk file.
Enter has two options: **Toxicity** and **Exposure**. **Toxicity** allows the entry of toxicity data, **Exposure** allows the entry of exposure data. Although the procedures of entering each category are identical, except that **Toxicity** asks for a Toxic substance name, while **Exposure** asks for an Exposure substance name, both sets are treated very differently inside *E_TX*. The toxicity data set is 'extrapolated' to HC_5 estimates, e.g. for setting standards, while exposure values are evaluated through their estimated hazard to species, on the basis of the toxicity data.
Enter is rather primitive, working on a line by line basis. It is impossible to go back to previous lines, in order to change them, or to delete previous entries. Typing errors, however, can be corrected through **Edit** (see **Edit**), after

completing data entry of the whole set. At present, one cannot add data to an already completed set, from within E_TX, nor can one delete one or more entries. This might change in future releases.
More flexible, full-screen editing, can be accomplished, however, with the aid of an editor or wordprocessor outside of E_TX. One has to save the data from E_TX, then, in order to apply these editors to the data.
`Enter` first asks for a toxic substance name, or exposure substance name. A highlighted entry line signifies that the Line edit mode is on (full highlight). Line edit mode is pretty sophisticated, though. One can enter a substance name, and edit the line with the Edit keys, until satisfied. Pressing **`[Esc]`** blanks the whole entry line. **`[Enter]`** terminates Line edit mode. If **`[Enter]`** is pressed immediately, without entering a name, **`Unnamed substance`** is assumed.
There is no check on the validity of a name entered, from the point of view of E_TX file reading conventions (see Files), so, if you do not follow these conventions, **`Enter`** may continue without complaint, the name may even be correctly saved to a file through **`Save`**, but E_TX may not be able to read in the file successfully lateron. The convention is: do not start the name with a numeric 'head', followed by white space (spaces and/or tabs). A numeric head, immediately followed by other non-white characters is fine (e.g. **`2,4-dimethyl ...`**). The reason is that E_TX will interpret the string with the white space as a concentration value followed by a comment, and assumes that the substance name was not given (**`Unnamed substance`**).
After successfully entering the substance name, one is prompted for the first numerical entry through **`1:`** followed by a full highlight, which means that Line edit mode is on again. This Line edit mode is exacty the same as the one for the toxic substance. Now, a numeric non-zero, non-negative, value, e.g. a concentration value, is expected. E_TX will respond with **`Entry xx is invalid`**, if it doesn't like it, and after pressing the **`[Enter]`** key, the entry line is blanked, and restored in full highlight. **`[Esc]`** blanks a non-empty entry line. If the entry is found OK, the next entry is prompted with **`2:`**. Up to a maximum of 300 data values can be entered for each data set.
One may enter bare numbers, and E_TX will not complain. But, one can make each entry self-documenting by adding additional information about the nature of the entry, such as the unit of measurement, a species name, a reference, anything informative. This information may be entered directly following the numeric value, i.e. in Edit line mode, but separated from it by at least one space or tab. The value and its comment string, as it is called, stay together as one data record. When the data is saved to a disk file, the complete data records are saved on a line by line basis.
To stop the entering process, press the **`[Enter]`** key while in Edit line mode,

with nothing entered on the highlight. E_TX will ask for confirmation, one may resume entering mode by pressing the **[Esc]** key. Pressing **[Esc]**, while in Edit line mode, with an *empty* highlight, will cause E_TX to ask you, whether you want to leave **Enter**ing mode, and restore the old data. If you confirm this, everything entered so far is deleted, and the previous situation (perhaps without data) is restored. To clear previous data, start **ETX/ Data/ Enter**, press **[Enter]** at the first numerical (empty) prompt, and press **[Enter]** to confirm. After completing **Enter**, E_TX immediately starts calculating all statistics and extrapolation estimates, as well as hazard estimates, if exposure data are available. Hence, there are no menu options or paths, that by themselves trigger calculations to be done. They have been done, as soon as data have been entered, edited, or read.
After having entered a complete data set, either toxicity data, or exposure data, or both, one can do several things. One can inspect the data, and browse through them, with **Show**. **Show** only lets you look through the data, while pointing (highlighting) individual entries. These are not full highlights, so you are not in Line edit mode. One can also **Edit** the data, that is make modifications, add comments and so on. It is impossible, right now, to add extra data to the sets, or to delete data. This has to be done outside of E_TX, with the aid of an ASCII editor, a wordprocessor, or a spreadsheet. A third possibility is to **Save** the data just entered to a disk file, leading to a permanent storage of the data, e.g. for distribution to colleagues, for coordinated data management within a work group, for outside editing, and so on. A fourth possibility is to go straight on to **Results/ Save**, and save the results of the calculations. A **Results** file does contain the data entered or read, so if you have a results file, the data can always be recovered. This is quite easy in fact, because the initial section of a results file is similar to a stand-alone E_TX toxicity or exposure data file. And of course, if you are just experimenting, you can walk the menu tree to study the results, and quit E_TX as soon as you are done. Then the data can not be recovered afterwards.

9. Exposure

Paths: **ETX/ Data/ Enter/ Exposure.**
ETX/ Data/ Read/ Exposure.
ETX/ Data/ Show/ Exposure/
ETX/ Data/ Save/ Exposure/
ETX/ Statistics/ Logarithms/ Exposure/

Exposure always refers to an exposure data set, comprising similar concentration

values relating to one particular toxic substance under study. These values may be environmental or experimental concentrations of that substance, either measured, estimated, proposed, legally imposed, or whatever. The purpose of the exposure data is, that, given an extrapolation exercise, that is, given an impression of the distribution of species sensitivities for a specific toxic substance, one would like to assess the percentage of species that is harmed, or is safe, with respect to a set of given exposure concentrations relating to the toxic substance under study. These may be exposures in the past, at present, to be expected in the future. They may come from a survey, a set of scenario predictions, etc. What E_TX does is to try to estimate percentiles of the species sensitivity distribution at the given exposure concentrations. These need not be raw data, they may be calculated from other data, e.g to make them comparable to the nature of the toxicity data. One may also massage the toxicity data first to make them comparable to the environmental data. That is pretty much the responsibility of the user.
E_TX can have one exposure data set of maximally 300 values, including their comments. One may **Enter** them, **Read** them from a disk file, browse through them after entering or reading (**Show**), **Save** them to a disk file after entering or reading, study their log transformed values (**Logarithms**) on which all statistics are based. The reader is referred to these options.
Exposure data cannot be **Edit**ed at present from within E_TX. If the exposure data need to be edited, then save them, and use an outside editor to make modifications. There is a work-around to do it inside E_TX, by retreiving them as toxicity data (see **Edit**), but this is only feasible if you understand the structure of E_TX files.

10. Extrapolation

Path: **ETX/ Extrapolation/**

Extrapolation is the third option of the E_TX main menu screen. Extrapolation, in E_TX, is the estimation of the hazardous concentration at the 5th percentile of the distribution of species sensitivities for a toxic substance, and forms the main focus of the program.
Activating **Extrapolation** leads to a submenu with three options, since there are three statistical distributions involved: logistic (**Logistic**), normal (**Normal**), and triangular (**Triangular**). However, the statistical treatment of the triangular distribution differs in several respects from those of the logistic and normal distributions. See **Triangular** for some extra information about that.

The logistic and normal distribution version of extrapolation are very much akin. In both cases, the screen reports the species protection level (95%), that cannot be changed, the number of toxicity tests involved (minimum number is two), and the Median Estimate of the HC_5, printed bold on the screen, as well as the 95% Underestimate of the HC_5. These are based on extrapolation constants, reported under **`Basic Statistics`**. Tables of the extrapolation constants are reproduced in Appendix B.
In general, there is very little difference between the extrapolation answers of the respective distributions. Note that the HC_5 estimates are in the *original units* of the data. Hence, the $\log_{10}$ transformation of the data, that had been applied to do the statistics, has been removed in the HC_5 estimates. These estimates can be either directly used for setting environmental standards or objectives, or, depending on other considerations of a scientific, or policy nature, further calculations, e.g. involving extra safety factors, or partition coefficients, etc., may be applied to arrive at the final answers wanted. These considerations are outside the scope of *E_TX*.

11. Files

E_TX data files can come into existence in essentially two ways. One is by saving data from inside *E_TX* to a file. The menu path is: **`ETX/ Data/ Save/ Toxicity/ As Is`**, to save *E_TX* toxicity data, and: **`ETX/ Data/ Save/ Exposure/ As Is`**, in order to save *E_TX* exposure data.
The second way of coming into existence of an *E_TX* data file is by constructing it yourself with an editor, word processor or spreadsheet. Both possibilities are treated in the next two paragraphs.
In the current version of *E_TX*, files are saved into the same directory as where **`ETX.EXE`** is located. One can only read files from this same directory. Hence, do not try to change directory from within *E_TX*. This might change in future versions.

a. *E_TX* generated files

Files saved by *E_TX* can be read in again by *E_TX*. Saving a data file through *E_TX*, has the advantage that *E_TX* automatically includes all kinds of information in the header of the file, to be shown below. This identifies the file internally, e.g. date, time, a save serial number, etc.. Since they are readable ASCII files, this information can be inspected, printed, and edited.
The advantage of self-constructed files is that you can include a lot more

additional information into the data file than is currently possible through E_TX: how the data were selected, scope of the project, extra comments between data lines to identify taxa, or circumstances, and so on.
E_TX data files can be made almost self-documenting. There is no excuse anymore for unannotated toxicity data. Of course, both ways of constructing E_TX files can be combined. For example, one may enter data interactively in raw order, then save them in sorted form, and edit the file through adding additional annotation. If the very flexible rules of E_TX data formatting are maintained, the resulting file can be read again by E_TX. When more data become available, they may be added to the file, with ample space for documentation.
E_TX data files can have any extension. However, toxicity data saved by E_TX have default extension **.TOX**, while exposure data have default extension **.EXP**. Existing data files can never be overwritten by E_TX.
Both toxicity and exposure data files obey the same rules of formatting. E_TX cannot make a distinction between them, so reading an exposure file into the toxicity data structure is possible, but likely to result into non-sensical calculations.
Table B3 displays a toxicity data file as saved by E_TX. In fact, the lines starting with a **!** have been added by E_TX automatically. These are comment lines and serve as comments. They tell us that it is an E_TX toxicity data set, that it was saved with *E_TX 1.3a*, the name of the file (chosen by the user), date, time, a data save serial number as an extra lable in case of trouble, the name of the toxic substance and additional information about it, the number of toxicity data, and then the seven data values.
Note, that these are annotated data. The annotated data strings were entered from the keyboard in E_TX, including the lay-out to align decimal points and unit names. The comment strings following the data may have any form, and may contain any information. We have copied the information from Table 2 in Van Straalen and Denneman (1989). Unit, species, data reference, special circumstances: everything fitting in one line (80 characters) may be included in the comment string. If you enter the data through E_TX, save through E_TX, and read back into E_TX, it is impossible to separate the data values and their comment strings, although the latter are optional. Moreover, from inside E_TX it is impossible to overwrite a file. It simply refuses to save data to an existing file.
Table B4 displays an exposure data file as saved by E_TX. It has the same basic structure as an toxicity data file. The exposure substance is named differently here. The environmental or exposure concentrations may be of a different nature than the laboratory data.

b. Editor generated files

Editor files are data files meant to be read by E_TX, but not made through E_TX. The advantage, as already mentioned, is flexibility in documenting the data sets. Moreover, data files can be managed by a data manager independently of E_TX. Any ASCII editor can be used, e.g. MS-DOS 5.0 Editor, Turbo Pascal Editor, etc. Also, wordprocessors, spreadsheet, and data base programs. Do not use the internal format of wordprocessors. Use there ASCII export options.
Table B5 shows a fancy E_TX data file composed with WordPerfect. The lines are saved as 'ASCII Text (DOS)'. In this case, we have not tried to mimic the E_TX data file header, although one can take one from an E_TX data file and adapt it. The purpose of this example is to show how flexible E_TX data formatting is. Blank lines (lines with spaces, tabs, and a Carriage Return), and Comment lines with an exclamation mark in front, perhaps preceded by white space (spaces and/or tabs), can be put at any place in the file: at the beginning, at the end, and anywhere in between.
The first item to alert E_TX, while reading a data file, is the toxic substance name, or exposure substance name. If one lacks, E_TX assumes **`Unnamed Substance`**. These names are not preceded be an exclamation mark, and may contain spaces, and any additional information fitting on the line. Substance names may start with numeric information, but *not followed by spaces or tabs*. This is because a number followed by white space and additional characters is interpreted by E_TX as data: a value followed by a comment. So, as a substance name: **`2,4,5-T`** is fine, but **`2 4,5-T`** would be read as the value **`2`** and comment **`4,5-T`**. After a number has been seen, no names without exclamation marks may follow anymore. They are considered erroneous and E_TX will complain about an error in line number xx.
Hence, E_TX data, e.g. annotated concentrations of toxicity tests, consist of a number, optionally followed by a comment string, separated by white space (blanks and/or tabs). One blank suffices. Note that at least one blank, or tab, is obligatory now. So, **`15mg/l`** is *not* OK, in order to prevent typographic errors to pass by unnoticed, e.g. **`1o.7 ug/l`**. E_TX counts the data automatically. It can correctly distill a few numbers hidden within a heavily commented file.
Separate data must be on separate lines: the leading number on a line is taken as the value, while what follows, after white space, is interpreted as comment string. So, the line: **`2.0 1.0 5.0 ug/l (three reproduction values for Daphnia)`** contains *one* E_TX data value (**`2.0`**), and its comment string: **`1.0 5.0 ug/l (three reproduction values for Daphnia)`**.
There is only one substance per file. One cannot use a subset of the numbers in a file. If that is wanted, use an editor to select the numbers and save them in a

new file.
Since both the substance name, as well as the data comment strings are optional, a minimum *E_TX* data file consists of a bare list of anonymous numbers, separated by Carriage Returns. This is perhaps useful for numerical experimentation.

c. *E_TX* Results files

With the option **`ETX/ Results/ Save/ All.`**, one can save the results of an *E_TX* session to a disk file. The results saved always refer to the latest run. Hence, if one enters some toxicity data, and later reads new ones from a file, all internal results relate to the latest run. The single option **`All`** indicates that every result, whether it has been on the screen or not is saved. One cannot save part of them, e.g. just the extrapolation results.
In fact, there is no need to study any result on the screen. One can enter or read in the data, and go straight on to **`ETX/ Results/ Save/ All.`**, to save a printable ASCII file with everything in it, including the data themselves.
Table B6 shows the contents of the *E_TX* results file **`CDVSTRAA.ETX`** that corresponds to the *E_TX* data files **`CDVSTRAA.TOX`** (Table B3) and **`CDVSTRAA.EXP`** (Table B4). These files are also on the distribution disk.
The Results file is an ASCII printable file and gives an overview of all possible screen output appended to each other, with some additional comments. One can import this file into an editor or word processor, edit the text, and print it as a whole or in parts.
Results files have default extension **`.ETX`**. This can be changed, when prompted, but giving no extension is overruled by *E_TX* by adding **`.ETX`**. *E_TX* cannot overwrite existing files.

12. Fixed Hazard

Path: **`ETX/ Hazard/ Logistic/ Fixed Hazard.`**

`Fixed Hazard` is the first option of **`ETX/ Hazard/ Logistic/`**. The second option is **`At Exposure`** (see there, and see **`Hazard`**, as well as Appendix A for the full story).
`Fixed Hazard` refers to the display of estimated hypothetical exposure values at a range of fixed hazard percentages (1%, 2%, 5%, 10%, 25%, 50%, 75%, 90%, 95%, 98%, and 99%). This yields an impression of what the range of concentration values is that gives rise to these hazards. The median estimate of

the HC_5 is printed bold on the screen.
One may also print this list (from the **`Results .ETX`** file) for a given toxic substance of special interest, and use it as a lookup table for future exposure concentrations.

13. Goodness-of-Fit

Path: **`ETX/ Statistics/ Goodness-of-Fit/`**

When fitting distributions to data, in order to estimate percentiles, and the HC_5, it is important to assess, whether the data indeed seem to derive from the hypothesized distribution. This can be done with a goodness-of-fit test. One may be familiar with the well-known Kolmogorov-Smirnov test for goodness-of-fit (see almost any textbook on statistics). However, this test is designed for situations where the distribution generating the data is known, as well as its parameter values. In our case, the distribution parameters are estimated from the data. Hence, we arrive at a problem of circularity: the distribution is estimated from the data, can we test whether the data derive from the distribution?
D'Agostino & Stephens (1986) show that one can approach this matter with the same Kolmogorov-Smirnov test statistic:

$$D \cdot \sqrt{n}$$

only with modified critical values for this statistic, given a pre-set significance level. *D* signifies the maximum distance between the empirical distribution function (staircase) of the data, and the estimated distribution. Their tests, especially for the logistic distribution, distinguishes between different sample sizes: 5, 10, 20, 50, and infinity. (D'Agostino & Stephens, 1986, p.158), which seems quite relevant to the usual size of toxicity data sets. The table is reproduced as Table B7 in Appendix B. The parameters of the logistic distribution must be estimated from the data by maximum likelihood. Thus, we had to estimate the parameters this way (see **`Basic Statistics`**).
For the smallest sample size feasible for an extrapolation exercise, $n = 2$, we found that for any values of the two data points, the test statistic equals 0.458. It seems impossible to derive critical values in that case. We reasoned that the interpolating curve of critical values for intermediate sample sizes should all intersect at 0.458 at sample size two. Hence, we added this case to the table of critical values (Table B7, Appendix B).
Intermediate critical values are derived in E_TX through linear interpolation. E_TX

reports the goodness-of-fit for all four significance levels, and prints whether the hypothesis that the data derive from the logistic should be rejected, or not. The choice of significance level is up to the user.
For the normal distribution, D'Agostino & Stephens (1986, p. 123), present a modified Kolmogorov-Smirnov test statistic:

$$D \cdot (\sqrt{n} - 0.01 + 0.85 / \sqrt{n})$$

Critical values of this statistic for the same four significance levels are reproduced in Table B8 of Appendix B. Now, no further differentiation for low sample size is presented. We are unclear about the validity of the test for sample sizes much below 20. Hence, *E_TX* does present the goodness-of -fit for the normal distribution for small sample sizes, but the warning is printed: **`Below n = 20, this test may not perform well`**.'

14. Hazard

Path: **`ETX/ Hazard/`**

`Hazard` is the fourth option of the *E_TX* main menu. Hazard assessment, as defined in *E_TX*, is discussed in Appendix A. The purpose of hazard assessment is to estimate the hazard to species at given environmental exposure concentrations, that may be independent of the toxicity data. The exposure data set may refer to any predicted or measured data, perhaps adapted for comparison to the laboratory toxicity data. The only option of **`Hazard`**, currently, is **`Logistic`**, since the primary emphasis of the HC_5 has been on the logistic distribution.
The density estimate of the logistic distribution employed is calculated in such a way that, if an exposure value happens to be equal to the median estimate of the HC_5 on the basis of the current toxicity data set, then the estimated hazard is equal to 5%. Clearly, exposure values below the estimated HC_5 lead to hazards smaller than 5%, while those above lead to larger hazard percentages.
Under **`Logistic`**, two options are offered: **`Fixed Hazard`** and **`At Exposure`**. The first option displays hypothetical exposure values at a range of fixed hazard percentages (1%, 2%, 5%, 10%, 25%, 50%, 75%, 90%, 95%, 98%, and 99%). This yields an impression of what the range of concentration values is that gives rise to these hazards. One may also print this list (from the **`Results .ETX`** file) for a given toxic substance of special interest, and use it as a lookup table for future exposure concentrations.

The second option takes the exposure data as entered, or read from file, and evaluates the hazard at these values.

15. Leave E_TX

Path: **ETX/ Quit/ Leave ETX!**

After pressing **Quit**, you can either change your mind, through the first option (default): **Main Menu**. But, if you are definitely sure to leave E_TX, you can do so by activating the second option of **Quit**, i.e. **Leave ETX**. The trailing exclamation mark in a permanent warning that all internal data, or results, are lost by doing so! This warning is independent of whether you have saved the data, or the results, or not. If you haven't, no extra warning, or prompt follows, so, take care, when leaving E_TX.

16. Logarithms

Path: **ETX/ Statistics/ Logarithms/**

Logarithms is the first option of **Statistics**. E_TX data are entered or read as raw concentration values, densities, accumulated amounts, or other toxic substance measures. All statistical calculations and percentile estimates are carried out with the $\log_{10}$ transformed data. One can examine the data converted to logarithms through **Logarithms**. Further options allow one to choose between **Toxicity** and **Exposure**.
The data are always log-transformed, this cannot be circumvented once inside E_TX. If log transformation is not wanted, one may construct a raw data set consisting of ani-logs, i.e. powers to the base 10. The logs are always taken with respect to the base 10. Log transformation requires that the raw data are positive (non-zero, non-negative). Non-positive concentrations, while entered or read, lead to **Invalid Entry** errors. If the original data do contain zero values, correct them before feeding them to E_TX, by adding a small value, e.g. half the detection limit. One may add this 'starter' value to all the data if wished. Do not forget to substract these small amounts from the extrapolation estimates afterwards. These estimates are in the original units.
If **Logarithms** is activated, and the appropriate data set chosen, the screen displays log concentrations followed by an arrow (**<--**), followed by the original data, including their comments. One may browse through the log values in the

original order (**As Is**), or in sorted order (**Sorted**), by using the arrow keys **[Up]**, **[Down]**, **[Home]**, and **[End]**. If there are more than 20 values, one can page through the data set with **[PgUp]**, **[PgDn]**, and/or **[Up]** and **[Down]**. The entry highlight moves with the arrow and paging keys. It only indicates where one in located in the set, no editing is possible. Consecutive entries are numbered with a number followed by a colon (**:**). The end of the data set is indicated with: **(no more)**. End of page, with more values to follow, is indicated with: **(more)**.

17. Logistic

Paths: **ETX/ Statistics/ Basic Statistics/ Toxicity/ Logistic.**
ETX/ Statistics/ Goodness-of-Fit/ Toxicity/ Logistic.
ETX/ Extrapolation/ Logistic.
ETX/ Hazard/ Logistic/

Logistic refers to the so-called logistic distribution. This is a statistical distribution of a certain mathematical form. See Aldenberg & Slob (1993) for a review of some logistic mathematics. The logistic distribution looks very much like the normal distribution. It has two parameters α and β, closely related to the mean and standard deviation of the distribution.
Basic Statistics displays some sample statistics of the toxicity data and different estimates of α and β.
Goodness-of-Fit shows measures of departure of the toxicity data from the logistic distribution. If these measures are too high, doubt is thrown on the hypothesis that the logarithms of the toxicity data derive from the logistic distribution.
Extrapolation treats the estimation of the hazardous concentration (HC_5), being the 5th percentile of the logistic distribution, from the current toxicity data. Extrapolation can also be based on the normal and on the triangular distribution.
Hazard handles the assessment of the percentage of potentially harmed species at the current exposure data set. Hazard assessment is only done for the logistic distribution in the current version. See the options just mentioned for more details.

18. Main Menu

Path: **ETX/ Quit/ Main Menu.**

If you want to quit *E_TX*, but change your mind, e.g. to see if the data were saved, you can return to the *E_TX* main menu, by activating **Main Menu**. The alternative option is **Leave ETX**, see there.
[F10] brings you up to the main menu from anywhere in the menu tree.

19. Normal

Paths: **ETX/ Statistics/ Basic Statistics/ Toxicity/ Normal.**
ETX/ Statistics/ Goodness-of-Fit/ Toxicity/ Normal.
ETX/ Extrapolation/ Normal.

Normal refers to the well-known normal distribution from ordinary statistics. The normal distribution has two parameters μ and σ, called mean and standard deviation.
Basic Statistics displays some sample statistics of the toxicity data and different estimates of μ and σ.
Goodness-of-Fit shows measures of departure of the toxicity data from the normal distribution. If these measures are too high, doubt is thrown on the hypothesis that the logarithms of the toxicity data derive from the normal distribution.
Extrapolation treats the estimation of the hazardous concentration (HC_5), being the 5th percentile of the normal distribution, from the current toxicity data. Extrapolation can also be based on the logistic and on the triangular distribution. No hazard assessment is incorporated in *E_TX* as yet for the normal distribution. See the options just mentioned for more information.

20. Printer

There is currently no direct printer support in *E_TX* (version 1.3 a). However, all results can be saved into an ASCII printable file. This file can be imported into an ASCII Editor, or a wordprocessor for editing and printing. See **Results**, or **All**.

21. Quit

Path: **ETX/ Quit/**

Through **Quit** one can leave E_TX. E_TX won't let you do so immediately. A submenu follows with the first option: **Main Menu**, as default, meaning: return to the E_TX main menu screen. The second option **Leave ETX!** does end E_TX, without any further delay. The exclamation mark is a reminder that no warning is given, if you haven't saved your data or results. Any data entered, or results obtained, but not saved, is lost after activating **Leave ETX**.
A quick way to get out, from anywhere in the menu tree, is **[Ctrl]+[End]**, and **[Ctrl]+[Y]** to confirm. This certainly destroys all data entered, or results obtained. An emergency stop is given by **[Ctrl]+[Break]**.

22. Read

Path: **ETX/ Data/ Read/**

Read is the second option of **ETX/ Data/**. With **Read**, one can read data into E_TX from disk files. E_TX data files are ASCII printable files that are either saved through E_TX, or files made trough an ASCII Editor, a wordprocessor, or a spreadsheet. Use: save (export) as ASCII, or as a print file (often extension **.PRN**), in these programs.
The E_TX file conventions are explained under Files of the Reference List.
Read has two options: **Toxicity** and **Exposure**. A data file read under **Toxicity** fills the E_TX internal toxicity data set. A data file read through **Exposure** is put into the E_TX internal exposure data set. There is no essential difference between a toxicity and an exposure data file. Both have the same structure. While saving, default extensions **.TOX**, and **.EXP** are standard, but not obligatory, E_TX conventions. One may develop other conventions. Additional internal comment lines in data files, saved by E_TX, further indicate, whether we have a toxicity data file or an exposure data file, but this information is not interpreted by E_TX. Hence, there is good reason to stick to some extension convention, preferably that of E_TX, next to the informative and critical annotation of the toxic or exposure substance names and of the individual data entries.
Here we assume that there are valid E_TX files available in the directory where **ETX.EXE** is located. After activating **Read**, and choosing the type of data set, the menu screen clears and E_TX responds with:

Edit Search Profile for Data File to be Opened:
***.TOX**

This means that the MS-DOS search profile conventions, including wild cards, are supported. The highlighted edit line indicates that Line edit mode is on, with ***.TOX** already filled in as a default search profile. It means: list all the files with extension **.TOX**. Pressing **[Enter]**, accepts the default search profile, and indeed lists the files. With the Edit keys, one may change the search profile, or fill in a filename of specific interest. When no file conforms to the user-supplied search profile, *E_TX* says: **No file found on this Search Profile**.
It is currently *not* possible to specify a file including its directory path, such as **MYDATA*.TOX** (for files in the subdirectory of the one where *E_TX* is located), or **C:\RIVM\ALD*.TOX** (for complete paths). This feature is high on the list to be changed in a future version.
Once you have a list of files on the screen, you can browse through them with the arrow keys. The highlight indicates the position of the list you are pointing at. After having found the file to load, you can load it by pressing **[Enter]**. *E_TX* tries to read the file, as explained under Files, and if it runs into an offending line, if any, it will beep and say: **File Reading Terminated: Invalid Input at Line *xx*** and then, print the invalid line. The file is not processed any further. See Files for the expectations *E_TX* has about the structure of the information in *E_TX* data files.
Otherwise, if no errors are found, the file is read, and the data are stored internally. The screen says: **File Reading Successful: *n* Data Read**, pauses for a while, because *E_TX* will now calculate all statistics immediately. This may take some time for a large data set.

23. Results

Path: **ETX/ Results/**

Results is the fifth option of the main *E_TX* menu screen. In the present version of *E_TX*, **Results** is only used for saving the results of the calculations to a disk file. Hence, the only option is **Save**. Other options may be added in future versions, e.g. printer support, results selection, and so on. **Save** in turn has the only option **All**, to indicate that it is only possible now to save all results relating to the current data sets (toxicity data and exposure data).
Collecting results is only necessary, if one wants to save the data and the results in one file, for documentation, or later printing. There is no control over what is saved, and what not, nor about the order of the items saved. Everything that

E_TX calculates is saved, whether it *has been on-screen, or not*. Hence, a very quick way to do an extrapolation exercise, and hazard assessment, is to enter data, or load (read) a file, and save all results. The resulting **.ETX** file can be printed out, and studied. No menu tree walking, or on-screen displaying, is really necessary to obtain the results. As soon as the data are entered, or read, all calculations are done.

24. Save

Paths: **ETX/ Data/ Save/**
ETX/ Results/ Save/

ETX/ Data/ Save/ lets us save the data to disk, that previously have been entered, or read from disk files. In the current version, it is only possible to save files to the directory where the program runs.
E_TX may contain two data sets at once, differing in nature. The one set is formed by the (laboratory) toxicity data set; the other is the (environmental) exposure data set. Hence, **ETX/ Data/ Save/** has two further options: **Toxicity** and **Exposure**. The first option saves the toxicity data, the second the exposure data, but not before in each case the sorting order has been selected (**As Is**, or **Sorted**).
For saving either data set, a filename has to be chosen. A default extension, **.TOX**, or **.EXP** is given on an inverse edit line. One can use the Edit keys (see there), to enter/edit the filename, e.g. **CADMIUM.TOX**, **24DMP.TOX**, etc. The extension may be changed to something else, e.g. **.DAT**, or **.TX1**, etc., or it can be removed. But *E_TX* will add **.TOX**, or **.EXP**, as appropriate, if no extension is given.
E_TX is designed not to overwrite any existing file. If you try, *E_TX* will refuse, and say so. This is done with an eye to current trends in Good Laboratory Practice. Of course, one can fiddle in the files through MS-DOS. In fact, one must do so to add, or remove, data to an *existing* data set. Hopefully, the lack of the possibility to overwrite files from inside *E_TX*, as well as the in-file information that is written, should catch most common accidents.
ETX/ Results/ Save/ refers to the writing of all statistical, extrapolation, and hazard assessment results, *including* the data, to a disk file. See **Results** and **All**. Here also, an inverse edit line is offered with a pre-written extension **.ETX**. One can use the edit keys to edit the filename, e.g. **CADMIUM.ETX**, etc. The extension may be changed, or removed. If no extension is given, *E_TX* will nevertheless add **.ETX** to the filename. Filenames longer than eight characters are trimmed. An existing file of the same name cannot be overwritten.

25. Show

Path: **ETX/ Data/ Show/**

Show is the third option of **Data**. After having entered data (**Enter**), or having read data from file (**Read**), one may want to inspect the data, in order to see what they look like, whether there are any typing or reading errors, whether the comment strings annotating the data have to be extended or modified, and so on (**Show**).
Show allows one to browse through both the toxicity and the exposure data sets. Hence, **Show** has two options: **Toxicity** and **Exposure**. One may further choose to browse the data in the original order (**As Is**), or in sorted order (**Sorted**). **Sorted** only refers to the on-screen representation. The data are kept in the original order internally.
One may browse through the data by using the arrow keys **[Up]**, **[Down]**, **[Home]**, and **[End]**. If there are more than 20 values, one can page through the data set with **[PgUp]**, **[PgDn]**, and/or **[Up]** and **[Down]**. The entry highlight moves with the arrow and paging keys. It only indicates where one in located in the set, no editing is possible. Consecutive entries are numbered with a number followed by a colon (**:**). The end of the data set is indicated with: **(no more)**. End of page, with more values to follow, is indicated with: **(more)**.

26. Sorted

Paths:
```
ETX/ Data/ Show/ Toxicity/ Sorted.
ETX/ Data/ Show/ Exposure/ Sorted.
ETX/ Data/ Save/ Toxicity/ Sorted.
ETX/ Data/ Save/ Exposure/ Sorted.
ETX/ Statistics/ Logarithms/ Toxicity/ Sorted.
ETX/ Statistics/ Logarithms/ Exposure/ Sorted.
ETX/ Hazard/ Logistic/ At Exposure/ Sorted.
```

Sorted always refers to the order of the data after sorting from smallest to largest. As can be seen from the above paths, **Sorted** refers to the last menu screen to fire the complete command. On this last menu, it is always the second option of two, the first being **As Is**, meaning unsorted. Hence, **As Is** is the default, while **Sorted** needs an extra key press (**[Down]**) before it can be activated with the **[Enter]** key. What **Sorted** actually does, depends on the previous menu options, to which the reader is referred.
An additional subtlety of **Sorted** is that, for the on-screen actions (all except

Save), no physical sorting is done: it only refers to the on-screen representation. The data themselves are kept in the original order. In case of **ETX/ Data/ Save/**, however, the respective data are really saved in sorted order to disk. If this option is used, without an analogous action through **As Is**, the original order of the data can never be recovered from the sorted files.

27. Statistics

Path: **ETX/ Statistics/**

Statistics is the second option of the E_TX main menu screen. It allows the user to look at some statistical summaries of the data. **Statistics** gives rise to a submenu. One can look through the logarithms of the data, which essentially enter the statistical analysis (**Logarithms**), inspect some basic statistical summaries of the data, like means, standard deviations, and distribution parameter estimates (**Basic Statistics**), and one can examine the goodness-of-fit of some statistical distributions fitting the data (**Goodness-of-Fit**).

28. Toxicity

Paths: **ETX/ Data/ Enter/ Toxicity.**
ETX/ Data/ Read/ Toxicity.
ETX/ Data/ Show/ Toxicity/
ETX/ Data/ Edit/ Toxicity.
ETX/ Data/ Save/ Toxicity/
ETX/ Statistics/ Logarithms/ Toxicity/
ETX/ Statistics/ Basic Statistics/ Toxicity/
ETX/ Statistics/ Goodness-of-Fit/ Toxicity/

Toxicity always refers to a toxicity data set, comprising similar concentration values relating to one particular toxic substance under study. These may be laboratory toxicity data for this toxic substance for different species or other taxa, either raw or averaged. A toxicity data set may also be a set of mesocosm, or field experiment, toxicity data for the particular toxic substance. The prime requirement is that it makes sense to assume that the toxicity data derive from a log-logistic, log-normal, or log-triangular distribution.
E_TX will try to estimate the 5th percentile for these distributions, based on the toxicity data set. This is called extrapolation. Furthermore, E_TX will try to estimate the distribution itself, in order to be able to estimate the hazard to

species, that is the percentage potentially harmed, at given environmental exposure concentrations. Hence, next to a toxicity data set, *E_TX* can have an exposure data set. Both sets refer to the same toxic substance, but their names (toxic substance name and exposure substance name) may differ, as well as their exact nature (total concentrations versus dissolved, or other fractions; conversion to standard conditions, e.g. sediment or soil characteristics).
Toxicity data may derive from a diversity of sources, as long as they refer to the same toxic substance. Some may be raw data, others calculated from other data, in order to obtain a consistent set, that is thought to derive from one statistical distribution. One may convert the toxicity data to values comparable to the exposure data, or adapt the exposure data to the laboratory conditions relating to the toxicity data. That is pretty much the responsibility of the user.
E_TX can have one toxicity data set of maximally 300 values, including their comments. One may enter them (**`Enter`**), read them from a disk file (**`Read`**), browse through them after entering or reading (**`Show`**), edit them to make modifications (**`Edit`**), or to extend the comment string (annotation); one may **`Save`** them to a disk file after entering or reading, study their log transformed values (**`Logarithms`**) on which all statistics are based.
One can study basic statistics of the toxicity data, as well as the distribution parameter estimates of distribution fitting them (**`Basic Statistics`**), and one can examine the goodness-of-fit of these distributions based on the toxicity data (**`Goodness-of-Fit`**) . The reader is referred to these options.

29. Triangular

Path: **`ETX/ Extrapolation/ Triangular.`**

`Triangular` is the third option of **`Extrapolation`**, and refers to the triangular distribution. The method implemented here is that of Erickson & Stephan (1988), called the FAV (Final Acute Value). See also OECD (1992). In the spirit of *E_TX*, the nature of the data is the responsibility of the user. Hence, the FAV may be applied to Chronic data, as well, and may be targeted on the taxon level preferred.
The statistics of the triangular distribution differs from the logistic and normal distributions. While the logistic and normal statistics are based on all the data points, the FAV is applied to the lower four points only. Since the four points are taken to derive from the lower tail of the distribution, they must be below the median of the data set. Hence, one needs a minimum of eight points to apply this method. The HC_5 is estimated from linear regression on these four points.

The advantage is that species insensitive to a toxic substance (the species at rank five or higher) have little influence on the estimate. On the other hand, the influence of the more sensitive species is increased.
One could device similar tail-oriented methods for the logistic and the normal distribution, and one could also develop all-points treatments for the triangular. We have not done so yet, since there is no need for a combinatorial proliferation of methods. In OECD (1992), it was found that the tail-oriented FAV on the basis of the triangular distribution yields results that correlate very well with the all-points median estimates of the HC_5 for the logistic and the normal distributions.
Perhaps, FAV falls into a caregory of its own. In a future release we might reorganize the **`Extrapolation`** section.

30. Version

The current program version is *E_TX* 1.3a. It is displayed on all *E_TX* menu screens as the first part of the menu title bar, e.g. **`ETX 1.3a/ Data/ Save/`**.

Appendix A Hazard Assessment

In this Appendix, our implementation of Hazard Assessment at given exposure concentrations is discussed, as well as some problems with a related formulation in the literature.

By Hazard Assessment in *E_TX*, we mean estimations of the hazard to species at one or more given concentrations, for example experimental or environmental concentrations. Such concentrations we call exposure concentrations.

For extrapolation, we work with the (laboratory) toxicity data at hand, and estimate the hazardous concentration for two levels of confidence. To estimate the hazard at concentrations that may or may not be equal to the hazardous concentration, or HC_5, we need an estimate of the statistical distribution that fits the toxicity data best.

Even if we confine ourselves to the logistic probability density function, there are several ways of fitting this distribution of species sensitivities to the toxicity data set. The first particular estimate, that comes to mind, is the density estimate based on the maximum likelihood estimate of the parameters of the logistic distribution. Another is based on the moment estimates of the parameters, and a third may be the one based also on the moment estimates, but now with the variance corrected for bias, i.e. $n/(n - 1)$ * variance.

All these estimates lead to different estimates of the logistic parameter β, and therefore different density estimates. These in turn result into different estimates of the hazard at given concentrations of exposure.

We decided that the density estimate should be the most accurate at concentrations around the HC_5, i.e. the hazardous concentration for 5% of the species.

In order to find the best β, we reasoned as follows. The 'best' estimate of the

HC_5 is taken here as the median estimate, that in the long run will overestimate the HC_5 as often as it underestimates it. Now the median estimate of the logHC5 is calculated as:

$$\log HC_5 = \bar{x} - k_L \cdot s_n$$

with k_L depending on sample size as tabulated in Appendix B (Table 1).
If we take

$$\alpha = \bar{x}$$

then we can calibrate β of the unknown density in such a way that the HC_5 of this calibrated density corresponds exactly to the median estimate of the true HC_5 on the basis of the sample. Thus we must solve beta from the identity

$$\alpha - \beta \cdot C_5 = \bar{x} - k_L \cdot s_n$$

where:

$$C_5 = \ln(\frac{95}{5}) = \ln(19) = 2.9444$$

See Aldenberg & Slob (1993) for a review of some logistic formulae.
This leads to the estimate

$$\beta = k_L \cdot s_n / C_5$$

Hence, β of the estimated density is proportional to the sample standard deviation.
For example let's take the seven Cd data (Van Straalen & Denneman 1989) again: 0.97, 3.33, 3.63, 13.5, 13.8, 18.7, 154 [ug/g]. We have (after

transformation with $\log_{10}$):

$$\bar{x} = 0.9712$$

and

$$s_7 = 0.7028$$

Hence, with

$$k_7\ (50\%) = 1.78$$

it follows that

$$\alpha = 0.9712$$

and

$$\beta = 1.78 * 0.7028 / 2.9444 = 0.4249$$

Now that we have estimated α and β, we can estimate percentages of hazard for species, *p*, at given $\log_{10}$-transformed concentrations:

$$x = \log_{10}(\textit{Exposure concentration})$$

This can be easily done from the explicit cumulative logistic distribution function:

$$p = \frac{100}{1 + \exp(-(x-\alpha)/\beta)}$$

So, the percentage of species *un*protected at a proposed reference value (Van

Straalen and Denneman (1989), Table 3) of 0.8 [ug/g], i.e.:

$$x = \log_{10}(0.8) = -0.0969$$

yields:

$$p = \frac{100}{1 + \exp(-(-0.0969 - 0.9712)/0.4249)} = 7.5\%$$

Similarly, at 3.5 [ug/g] (between 2nd and 3rd NOEC), we would calculate a species hazard of:

$$p = 26.8\%$$

At 15 [ug/g] (between 5th and 6th NOEC), the hazard is

$$p = 61.8\%$$

These hazard estimates seem in line with the scatter of the seven toxicity data. Van Straalen and Denneman, however, calculate 15% hazard for the case where we have 7.5% (Table 3: 85% protection). The explanation is that the Van Straalen and Denneman formula (9) is based on the 95% confidence estimate of the HC_5 instead of the median estimate used here. Writing their formula (9) as:

$$p = \frac{100}{1 + \exp((\alpha - x)/\beta)}$$

with:

$$\alpha = \bar{x}, \quad x = \log_{10}(c), \quad \beta = 3 \,.\, d_n \,.\, s_n \,/\, \pi^2$$

we observe that the formulae are identical, except that their beta is proportional to the sample standard deviation with a different constant of proportionality. The essential difference, apart from the mathematical form, is that the constant dm refers to the 95% 'confidence' column in Kooijman (1987, Table 1, delta2=0.05).

To illustrate, for the n = 7 case above, Van Straalen and Denneman have:

$$\beta = 3*2.82*0.7028/3.1416^2 = 0.6024$$

while we employ:

$$\beta = 1.78*0.7028/2.9444 = 0.4249$$

Indeed, by using the larger beta, we also arrive at p = 14.5%, i.e. 85.5% protection. Hence the difference between the two hazard estimates boils down to Van Straalen and Denneman having a broader density estimate on the basis of the same data. One would be inclined to think that hazard percentages sensu Van Straalen and Denneman at given concentrations tend to overestimate the true hazard percentages, and therefore, due to the 95% confidence factor d_n, could act as right confidence limits of these hazards. But, this has turned out not to be the case. If the log transformed exposure concentration happens to be equal to the sample average, both hazard percentages become 50%, while for exposure concentrations above the sample average, the Van Straalen and Denneman hazard percentages are *smaller* than those calculated on the basis of the median estimate. We think that an estimate of the species hazard at given exposure concentrations should not be based on an estimated standard deviation of the logistic distribution that results from matching the 5th percentile of that distribution with a deliberate underestimate of the true HC_5.

Appendix B Tables

Table B1. Extrapolation constants (Logistic distribution) for the 95% confidence underestimate and median estimate of the log HC_5.

n	95% confidence	median
2	27.70	2.49
3	8.14	2.05
4	5.49	1.92
5	4.47	1.85
6	3.93	1.81
7	3.59	1.78
8	3.37	1.76
9	3.19	1.75
10	3.06	1.73
11	2.96	1.72
12	2.87	1.72
13	2.80	1.71
14	2.74	1.70
15	2.68	1.70
20	2.49	1.68
30	2.28	1.66
50	2.10	1.65
100	1.95	1.64
200	1.85	1.63
500	1.76	1.63
inf	1.62	1.62

Table B2. Extrapolation constants (Normal distribution) for the 95% confidence underestimate and median estimate of the log HC_5.

n	95% confidence	median
2	26.206	2.35
3	7.656	1.94
4	5.144	1.82
5	4.210	1.78
6	3.711	1.77
7	3.401	1.76
8	3.188	1.74
9	3.032	1.72
10	2.911	1.70
11	2.815	1.69
12	2.736	1.68
13	2.670	1.68
14	2.614	1.68
15	2.566	1.68
20	2.396	1.67
30	2.220	1.67
50	2.065	1.67
100	1.927	1.65
200	1.840	1.65
500	1.763	1.645
inf	1.645	1.645

Table B3. Toxicity data set as saved by *E_TX* in file **CDVSTRAA.TOX**.

```
! ETX Data File: Toxicity Tests
! Saved with: ETX 1.3a
! Data File Name: cdvstraa.tox
! Date: Feb. 26, 1993
! Time: 15:46:42
! Data Save Serial Number: 1

! Toxic Substance:
Cadmium NOECs (Van Straalen & Denneman, 1989)

! Number of Data = 7

154    ug/g Dendrobaena rubida. Bengtsson et al. (1986)
 13.5  ug/g Lumbricus rubellus. Ma (1982)
 13.8  ug/g Eisenia foetida.    Malecki et al. (1982)
  3.63 ug/g Helix aspersa.      Russell et al. (1981)
  3.33 ug/g Porcellio scaber.   Van Capelleveen (1987)
  0.97 ug/g Platynothrus peltif.Van Straalen et al. (1989)
 18.7  ug/g Orchesella cincta.  Van Straalen et al. (1989)
```

Table B4. Exposure data set as saved by *E_TX* in file **CDVSTRAA.EXP**.

```
! ETX Data File: Exposure Values
! Saved with: ETX 1.3a
! Data File Name: cdvstraa.exp
! Date: Feb. 26, 1993
! Time: 15:47:27
! Data Save Serial Number: 2

! Exposure Substance:
Cadmium

! Number of Data = 1

  0.8 ug/g Reference Value proposed
```

Table B5. WordPerfect generated *E_TX* data file (**ASCII Text (DOS)**).

```
! This is an ETX data file

  ! describing the data set
  ! date, time, editor used, analyst, project
       !  any comments at any place

! The toxic substance:
  2,4-dimethylphenol

  0.1  ug/l Daphnia (geom. average of 3 reproduction tests)
  1.5  ug/l Pseudomonas (Canton, 1972b)
200    ug/l Lymnea (regression estimate)

! comments, empty lines  in between...
! (No Arthropods)

  0.67 ug/l Xenopus (Slooff, 1989)

! ...and at the end, plus some empty lines to follow
```

Table B6. Results as saved by *E_TX* in file **CDVSTRAA.ETX**.

```
! ETX Results File
! Saved with: ETX 1.3a
! Results File Name: CDVSTRAA.ETX
! Date: May  17, 1993
! Time: 18:45:14
! Results Save Serial Number: 1

! Toxic Substance:
Cadmium NOECs (Van Straalen & Denneman, 1989)

! Number of Data = 7
! Data Read from File CDVSTRAA.TOX

154    ug/g Dendrobaena rubida. Bengtsson et al. (1986)
 13.5  ug/g Lumbricus rubellus. Ma (1982)
 13.8  ug/g Eisenia foetida.    Malecki et al. (1982)
  3.63 ug/g Helix aspersa.      Russell et al. (1981)
  3.33 ug/g Porcellio scaber.   Van Capelleveen (1987)
  0.97 ug/g Platynothrus peltif.Van Straalen et al. (1989)
 18.7  ug/g Orchesella cincta.  Van Straalen et al. (1989)

---------Exposure Data--------------------

! Exposure Substance:
Cadmium

! Number of Data = 1
! Data Read from File CDVSTRAA.EXP

  0.8 ug/g Reference Value proposed

---------Log10 of Toxicity Data-----------------

As Entered/Read:

    Log10(Data)    Data
1:  2.1875 <-- 154    ug/g Dendrobaena rubida. Bengtsson et al. (1986)
2:  1.1303 <--  13.5  ug/g Lumbricus rubellus. Ma (1982)
3:  1.1399 <--  13.8  ug/g Eisenia foetida.    Malecki et al. (1982)
4:  0.5599 <--   3.63 ug/g Helix aspersa.      Russell et al. (1981)
5:  0.5224 <--   3.33 ug/g Porcellio scaber.   Van Capelleveen (1987)
6: -0.0132 <--   0.97 ug/g Platynothrus peltif.Van Straalen et al. (1989)
7:  1.2718 <--  18.7  ug/g Orchesella cincta.  Van Straalen et al. (1989)

Sorted:

    Log10(Data)    Data
1: -0.0132 <--   0.97 ug/g Platynothrus peltif.Van Straalen et al. (1989)
2:  0.5224 <--   3.33 ug/g Porcellio scaber.   Van Capelleveen (1987)
3:  0.5599 <--   3.63 ug/g Helix aspersa.      Russell et al. (1981)
4:  1.1303 <--  13.5  ug/g Lumbricus rubellus. Ma (1982)
5:  1.1399 <--  13.8  ug/g Eisenia foetida.    Malecki et al. (1982)
6:  1.2718 <--  18.7  ug/g Orchesella cincta.  Van Straalen et al. (1989)
```

```
7:  2.1875 <-- 154    ug/g Dendrobaena rubida. Bengtsson et al. (1986)

---------Log10 of Exposure Data-----------------

As Entered/Read:

    Log10(Data)    Data
  1:   -0.0969 <--   0.8 ug/g Reference Value proposed

Sorted:

    Log10(Data)    Data
  1:   -0.0969 <--   0.8 ug/g Reference Value proposed

---------LOGISTIC---------------------------------

Toxic Substance: Cadmium NOECs (Van Straalen & Denneman, 1989)

Sample statistics (log10s):
   XAverage =   0.9712
XStDev(n-1) =   0.7028
          n =   7
Extrapolation Constant (50%) =  1.7800
Extrapolation Constant (95%) =  3.5900

LOGISTIC Distribution Parameter Estimates (log10s):

(Moment Estimates, Bias Corrected)
   AlphaHat =   0.9712
    BetaHat =   0.3875

(Maximum Likelihood Estimates)
   Alphahat =   0.9445
    Betahat =   0.3727

(HC5 Fitting Estimates)
   AlphaHat =   0.9712
 MedBetaHat =   0.4248


Toxic Substance: Cadmium NOECs (Van Straalen & Denneman, 1989)

LOGISTIC Distribution, Goodness-of-Fit.

Number of Tests (n) = 7
Goodness-of-Fit, Kolm.-Smirn.: D*sqrt(n) =  0.512

Kolm.Smirn.   Crit.val.   Signif.   Logistic?
     0.512       0.657     10  %    Accepted
     0.512       0.699      5  %    Accepted
     0.512       0.743      2.5%    Accepted
     0.512       0.780      1  %    Accepted

---------Extrapolation, Logistic------------------

Toxic Substance: Cadmium NOECs (Van Straalen & Denneman, 1989)

LOGISTIC Distribution
Species Protection Level = 95%
```

```
Number of Tests = 7

Median Estimate Hazardous Concentration  =  5.2520E-0001
Underestimate (95% Confid.) Hazard.Conc. =  2.8076E-0002

---------Hazard Assessment, Logistic-------------

Toxic Substance: Cadmium NOECs (Van Straalen & Denneman, 1989)

Exposure Subst.: Cadmium

Exposure Values at Fixed Hazard Percentages:

Hazard      Exposure Value
   1%  <--  1.0448E-0001
   2%  <--  2.0789E-0001
   5%  <--  5.2520E-0001
  10%  <--  1.0909E+0000
  25%  <--  3.1953E+0000
  50%  <--  9.3593E+0000
  75%  <--  2.7414E+0001
  90%  <--  8.0299E+0001
  95%  <--  1.6679E+0002
  98%  <--  4.2135E+0002
  99%  <--  8.3836E+0002

---------Hazard at Exposure Data-----------------

As Entered/Read:

       Hazard       Exposure Data
  1:     7.49%  <--   0.8 ug/g Reference Value proposed

Sorted:

       Hazard       Exposure Data
  1:     7.49%  <--   0.8 ug/g Reference Value proposed


---------NORMAL----------------------------------

Toxic Substance: Cadmium NOECs (Van Straalen & Denneman, 1989)

Sample statistics (log10s):
   XAverage =   0.9712
XStDev(n-1) =   0.7028
          n =   7
Extrapolation Constant (50%) =  1.7600
Extrapolation Constant (95%) =  3.4010

NORMAL Distribution Parameter Estimates (log10s):

(Moment Estimates, Bias Corrected)
      MuHat =   0.9712
   SigmaHat =   0.7028

(Maximum Likelihood Estimates)
    MuHatML =   0.9712
 SigmaHatML =   0.6506
```

```
Toxic Substance: Cadmium NOECs (Van Straalen & Denneman, 1989)

NORMAL Distribution, Goodness-of-Fit.

Number of tests (n) = 7
Goodness-of-Fit, Kolm.-Smirn.: D*(sqrt(n)-0.01+0.85/sqrt(n)) =  0.566

Beware: Below n=20, this Test may NOT Perform Well.

Kolm.Smirn.   Crit.val.   Signif.    Normal?
     0.566       0.819     10  %    Accepted
     0.566       0.895      5  %    Accepted
     0.566       0.995      2.5%    Accepted
     0.566       1.035      1  %    Accepted

---------Extrapolation, Normal----------------------

Toxic Substance: Cadmium NOECs (Van Straalen & Denneman, 1989)

NORMAL Distribution
Species Protection Level = 95%
Number of Tests = 7

Median Estimate Hazardous Concentration  =  5.4248E-0001
Underestimate (95% Confid.) Hazard.Conc. =  3.8120E-0002

---------No Hazard Assessment, Normal---------------

---------TRIANGULAR---------------------------------

---------Extrapolation, Triangular.

Toxic Substance: Cadmium NOECs (Van Straalen & Denneman, 1989)

To calculate the Final (Acute) Value, the minimum number of tests
needed is 8. For toxic substance 'Cadmium NOECs (Van Straalen & Denneman,
1989)'
you have entered only 7 tests

---------End of ETX Results (All) file---------------
```

Table B7. Goodness-of-Fit for the Logistic distribution. Critical values of

$$D \cdot \sqrt{n}$$

at four levels of significance (adapted from D'Agostino & Stephens, 1986, p.158).

n	10%	5%	2.5%	1%
2	0.458	0.458	0.458	0.458
5	0.643	0.679	0.723	0.751
10	0.679	0.730	0.774	0.823
20	0.698	0.755	0.800	0.854
50	0.708	0.770	0.817	0.873
inf	0.715	0.780	0.827	0.886

Table B8. Goodness-of-Fit of the Normal distribution. Critical values of

$$D \cdot (\sqrt{n} - 0.01 + 0.85 / \sqrt{n})$$

at four levels of significance (D'Agostino & Stephens, 1986, p.123).

	10%	5%	2.5%	1%
n	0.819	0.895	0.995	1.035

References

Aldenberg, T., Slob, W. (1993) Confidence Limits for Hazardous Concentrations Based on Logistically Distributed NOEC Toxicity Data. Ecotox.Environm.Safety, 25, 48-63.

D'Agostino, R.B., Stephens, M.A. (1986) Goodness-of-fit techniques. Dekker.

Erickson, R.J., Stephan, C.E. (1985) Calculation of the Final Acute Value for Water Quality Criteria for Aquatic Organisms. National Technical Information Service, Springfield, Virginia, USA. NTIS No.: EPA 600/3/88/018. PB88/214994.

Kooijman, S.A.L.M. (1987) A safety factor for LC_{50} values allowing for differences in sensitivity among species. Wat. Res., 21, 269-276.

OECD (1992) Report of the OECD Workshop on the Extrapolation of Laboratory Aquatic Toxicity Data to the Real Environment. OECD Environment Monographs, No. 59. OECD, Paris.

Slooff, W. (1992) RIVM Guidance Document. Ecotoxicological Effect Assessment: Deriving Maximum Tolerable Concentrations (MTC) from single-species toxicity data. RIVM Report, 719102018. RIVM, Bilthoven.

Van Straalen, N.M., Denneman C.A.J. (1989) Ecotoxicological Evaluation of Soil Quality Criteria. Ecotox.Environm.Safety, 18, 241-251.

Wagner, C., Lokke, H. (1991) Estimation of Ecotoxicological Protection Levels from NOEC Toxicity Data. Water Res., 25, 1237-1242.

Index

RIVM report 601501028/2004

ETX 2.0

A Program to Calculate Hazardous Concentrations and Fraction Affected, Based on Normally Distributed Toxicity Data

P.L.A. van Vlaardingen*, T.P. Traas,
A.M. Wintersen and T. Aldenberg

This investigation has been performed for the account of the Directorate-General for Environmental Protection, Directorate for Chemicals, Waste and Radiation, in the context of the projects 'Setting (Inter)national Environmental Quality Standards'.

*Corresponding author. RIVM, Expert Centre for Substances, P.O. Box 1, 3720 BA Bilthoven, the Netherlands. E-mail: etx.info@rivm.nl

ETX 2.0 was programmed by Arjen Wintersen, working for Augustijn Onderzoek, WG-plein 406, 1054 SH Amsterdam.

National Institute for Public Health and the Environment, P.O. Box 1, 3720 BA Bilthoven, the Netherlands.
Telephone: 31-30-2749111; fax: 31-30-2742971

Rapport in het kort

E_TX 2.0. Een programma om 'hazardous concentrations' en 'fraction affected' te berekenen, gebaseerd op normaal verdeelde toxiciteitsgegevens.

Dit rapport is geschreven als handleiding bij het software programma *E_TX* 2.0. De rekentechnieken die met dit programma worden aangeboden worden onder andere gebruikt binnen het RIVM project '(Inter)nationale normstelling stoffen' (INS) maar ook bij de Europese risicobeoordeling van bestaande stoffen. In deze projecten worden milieurisicogrenzen afgeleid voor chemische stoffen. Zowel het INS- als het EU-raamwerk staat het gebruik van statistische extrapolatie toe wanneer voldoende toxiciteitsgegevens beschikbaar zijn. De resultaten van deze extrapolatie dienen als basis voor een milieurisicogrens (INS) of een geen-effect niveau (EU bestaande stoffen) van een chemische stof. In de wetenschappelijke literatuur is recent een methode beschreven om deze statistische berekening uit te voeren. Het programma *E_TX* 2.0 maakt deze methode toegankelijk voor hen die werkzaam zijn in de vakgebieden van de risicobeoordeling en/of normstelling. Het programma wordt samen met dit rapport verspreid op CD-ROM.

Trefwoorden: computerprogramma; soortsgevoeligheidsverdeling; statistische extrapolatie; risicobeoordeling.

Abstract

ETX 2.0. A Program to Calculate Hazardous Concentrations and Fraction Affected, Based on Normally Distributed Toxicity Data.

This report was written as a manual for the software program, *ETX* 2.0. The calculation techniques offered here are currently in use in the RIVM project, 'Setting (inter)national environmental quality criteria' (INS), and the EU risk assessment for existing substances. Environmental risk limits for chemical substances are derived here. Both the INS and EU frameworks allow for statistical extrapolation in the presence of sufficient toxicity data, which serves as a basis for an environmental risk limit (INS) or a predicted no-effect concentration (EU-existing substances) of a chemical substance. A statistical technique for achieving this result, recently described in the scientific literature, is made accessible to those working on risk assessment and/or standard-setting through the *ETX* 2.0 program. A CD-ROM is delivered along with the report.

Keywords: software; species sensitivity distribution; statistical extrapolation; risk assessment; hazardous concentration.

Contents

Samenvatting

Dit rapport is de handleiding van het softwareprogramma *ETX* 2.0. De rekentechnieken die dit programma biedt, zijn in gebruik binnen het RIVM project 'Internationale normstelling stoffen' (INS). Binnen dit project worden milieurisicogrenzen ('milieunormen') afgeleid in opdracht van het ministerie van VROM. Het richtsnoer voor de afleiding van deze risicogrenzen staat het gebruik van een statistische extrapolatietechniek toe, bij voldoende toxiciteitsgegevens. Het resultaat hiervan dient als basis voor een milieunorm. Een statistische techniek die voor dit doel geschikt is, is recentelijk beschreven in de wetenschappelijke literatuur. Het programma *ETX* 2.0 maakt de beschreven methode meer toegankelijk voor hen die werkzaam zijn op het gebied van normstelling. Het programma is ook bruikbaar bij risicoschattingen zoals die bijvoorbeeld worden uitgevoerd in Europese risicobeoordelingen van bestaande stoffen. Het programma wordt verspreid op CD-ROM, samen met dit rapport.

ETX 2.0 rekent op basis van de beschikbare toxiciteitsgegevens een zogenaamde soortsgevoeligheidsverdeling (SSD) uit. Vervolgens toetst het programma of deze verdeling voldoet aan de criteria van een normale verdeling. Van de SSD worden het 5^e percentiel en de mediaan berekend, beide met hun 90% betrouwbaarheidsinterval. Met de berekende SSD kan ook een fractie aangetaste soorten worden geschat bij een gegeven milieuconcentratie, of een verwacht ecologisch risico (EER) bij een serie van milieuconcentraties. Het programma biedt ook de mogelijkheid om van zeer kleine datasets het 5^e percentiel te schatten met behulp van de zogenaamde *small sample* methode.

Summary

This report is the manual of the software program *ETX* 2.0. The calculation techniques that are offered with this program are used within the RIVM project 'Setting international environmental quality criteria' (INS). Within this project, environmental risk limits ('environmental standards') are derived, by order of the Dutch Ministry of Housing, Spatial Planning and the Environment (VROM). The guidance followed for derivation of environmental risk limits allows for statistical extrapolation when sufficient toxicity data are available. The result of this extrapolation serves as a basis for an environmental standard. A statistical technique to achieve this result has recently been described in the scientific literature. The program *ETX* makes this method more accessible to those that work in the field of risk assessment of chemicals. *ETX* can also be used in effect assessments as carried out in e.g. the European risk assessments for existing substances. The program is distributed on CD-ROM together with this report.

ETX 2.0 calculates a normal distribution through the toxicity data entered by the user. This gives a so-called species sensitivity distribution (SSD). This distribution is subsequently tested on normality using statistical criteria. Of the calculated distribution, the estimated 5th percentile and median are presented, each with their respective two-sided 90% confidence interval. With the calculated SSD also the fraction of affected species at a given environmental concentration can be estimated, or an expected ecological risk (EER) at a series of environmental concentrations. The program also offers the opportunity to estimate the 5th percentile of very small data sets using the so-called 'small sample' method.

1. Introduction

This report is a user's manual to *ETX* 2.0. This program implements the theory of calculating hazardous concentrations and fraction affected from species sensitivity distributions as described in Aldenberg and Jaworska (2000). The program was developed in a stepwise manner, triggered by (i) the publication of the paper by Aldenberg and Jaworska (2000), (ii) the publication of the book 'Species sensitivity distributions in ecotoxicology' (Posthuma *et al.*, 2002) and (iii) the possibility to incorporate the 'small sample' method based on the papers of Luttik and Aldenberg (1997) and Aldenberg and Luttik (2002).

The first version of *ETX* was *ETX* 1.3a (Aldenberg, 1993) which runs under MS-DOS. *ETX* 1.3a enabled estimation of hazardous concentrations using logistic, normal and triangular distributions.

The sole purpose of *ETX* 2.0 is to offer the user the calculation methods as described in the supporting literature. The selection of data, applicability of the supporting theory to the data entered and interpretation of results is left entirely to the responsibility of the user.

ETX 2.0 is a software program written in Microsoft ® Visual Basic.NET (Edition 2003). Data can be entered simply as a column of values in the data input section. After performing the calculation, the output section shows statistics and graphical representations in different sheets.

Earlier versions of this program (in Microsoft ® Excel, called '*ETX* 2000') are now obsolete. All features and possibilities in earlier Excel versions are still present in *ETX* 2.0.

1.1 Referring to *ETX* 2.0

If you use the program in an official publication, book or report, it can be referred to it in the following way:

Van Vlaardingen PLA, Traas TP, Wintersen AM, Aldenberg T. 2004. *ETX* 2.0. A program to calculate hazardous concentrations and fraction affected, based on normally distributed toxicity data. Bilthoven, the Netherlands: National Institute for Public Health and the Environment (RIVM). Report no. 601501028/2004, 68 pp.

2. What can *ETX* 2.0 do for me?

As stated in the introduction, *ETX* 2.0 offers you the opportunity to apply statistical theory common to species sensitivity distributions (SSDs). This theory is most commonly applied in the field of ecotoxicology. In this scientific discipline, datasets will usually consist of end-points derived from toxicity studies with a given substance and a particular species or process representing a pre-defined environmental compartment or ecosystem. To speak in a more practical sense, your data may (e.g.) be chronic NOECs for freshwater organisms.

You can use *ETX* 2.0 to calculate the following items:

- The program calculates a normal distribution through your data set.
- The program will show the results of three goodness-of-fit tests that you can use to decide whether your data follow a normal distribution. The three tests are known as Kolmogorov-Smirnov, Anderson-Darling and Cramér-von Mises.
- Next, the program calculates the median HC_5 (hazardous concentration) plus its two-sided 90% confidence limit and the median HC_{50} plus its two-sided 90% confidence limit. At the HC_5 and HC_{50}, the corresponding median FA (fraction affected) is given (i.e. 5% and 50%, respectively) together with its two-sided 90% confidence limit.
- Results are graphically presented in a histogram and in a cumulative density function (the latter is commonly referred to as SSD).

If you have calculated a species sensitivity distribution, you have the possibility to calculate the FA at a given exposure concentration. The program gives you the median FA and its lower and upper estimates (5 and 95% confidence). If you have a series of exposure concentrations, you can also enter this series. This series is tested for normality of distribution with the Anderson-Darling goodness-of-fit test. Next, the expected ecological risk (EER) is calculated, making use of both the SSD and environmental concentrations. The result is graphically presented as a joint probability curve (JPC).

If you want to estimate the hazardous dose for 5% of mammals or birds (HD_5), but you only have a very small data set, you may want to use (what we call) the 'small sample' method. A standard deviation from a known, external data set of toxicity data is used to estimate the 5^{th} percentile of your data set. The lower and upper limit of the two-sided 90% confidence interval are also calculated.

Note that the 'small sample' method is dependent on knowledge of both the toxicity of your compound to a few species, as well as an expectation of the variation in toxicity data of that compound, derived from a high number of data (in the form of an external standard deviation). The method of deriving HC_5 and HC_{50} values or FA values can be applied generically to any set of toxicity data (n>1).

3. User's guide

3.1 Program type and system requirements

E_TX 2.0 is a Microsoft© Windows application that runs using the Microsoft©.NET (dotnet) framework. The program should therefore be run on Personal Computers (PC) equipped with an operating system that supports Microsoft applications. The Visual Basic.NET compact framework (version 1.1) is included on the CD and will be installed if necessary. Proper functioning of *E_TX* was tested on several combinations of Windows operating systems and Microsoft Office versions. We cannot guarantee proper functioning of *E_TX* under Windows 98. In most occasions, we encountered no problems, while in some cases installation under Windows 98 was not successful for unknown reasons. For later versions of Windows, installation was successful in most cases. In those cases where installation was unsuccessful, installation of a recent version of Microsoft's Data Access Components (which can also be found on your *E_TX*-CD) solved the problem. In Appendix 1, a list of combinations tested and possible solutions when installation problems are encountered, is given.

An operating system (like Windows NT or XP) is necessary. For proper functioning, *E_TX* does not need MS Office. However, to use the export option, MS Excel is needed, because export reports are generated in an MS Excel spreadsheet. However, the program works equally well without making use of this export option.

3.2 Installing *E_TX* 2.0

- Insert the *E_TX*-CD in the CD drive of your PC.
- Start the Windows Explorer (Start button, Programs, Accessories, Windows Explorer).
- In the left side of your screen, click on the drive that is your CD drive.
- In the right half of your screen, double click on the file **Setup.Exe**.
- The installation procedure will now start. The following screen may appear:

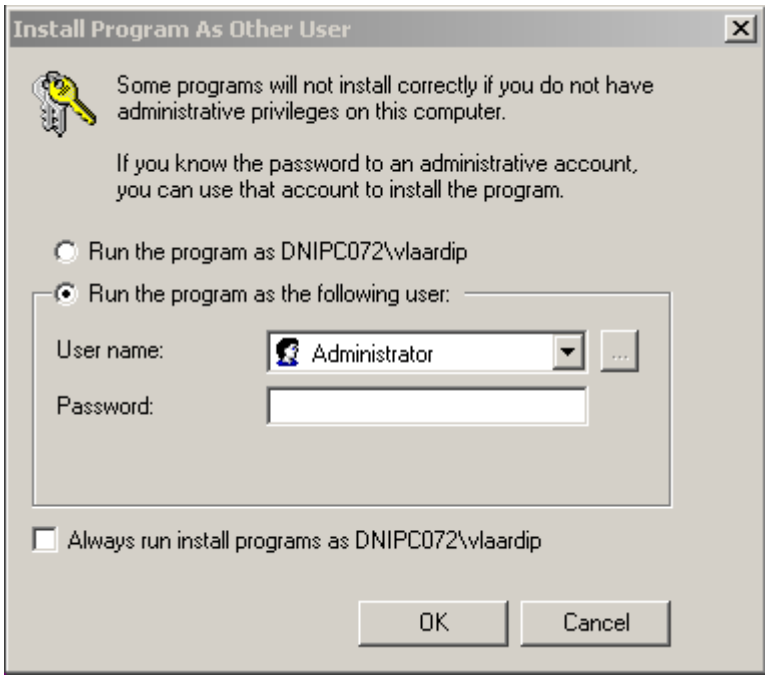


Figure 1. Intallation procedure – User screen.

You are asked to decide whether a person with administrator rights should install the program for you via the displayed dialog box. The choice for one of the options depends on your company's protocols. It is advised to install the program using a user account with administrator rights. Click **OK** after having selected the correct user.

– In the following screen, click **Next**.

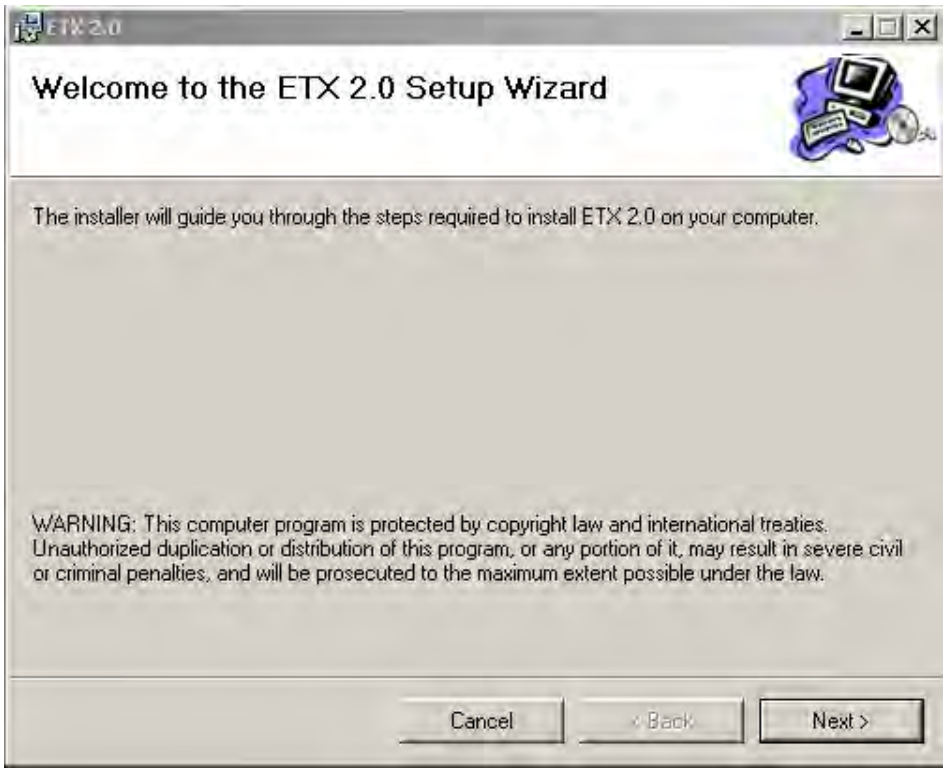


Figure 2. Installation procedure - Welcome screen.

– Carefully read the license agreement that is displayed in the following dialog box:

Figure 3. Installation procedure – License Agreement screen.

– If you agree with the content of this agreement, select the **I agree** option, and then click **Next**.

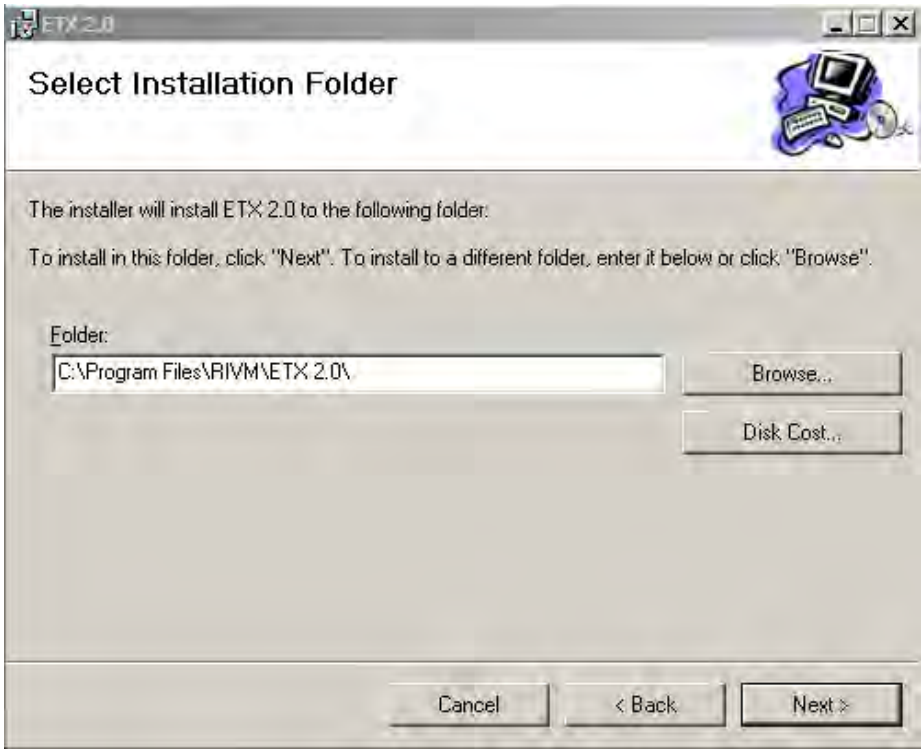


Figure 4. Installation procedure – Select Folder screen.

– In the dialog box displayed above, select the Folder in which you want E_TX to be installed.

The default directory is C:\Program Files\RIVM\ETX 2.0. If you want to install *ETX* in an other directory, click Browse and select the directory of your choice.

– Click on **Next** to continue.

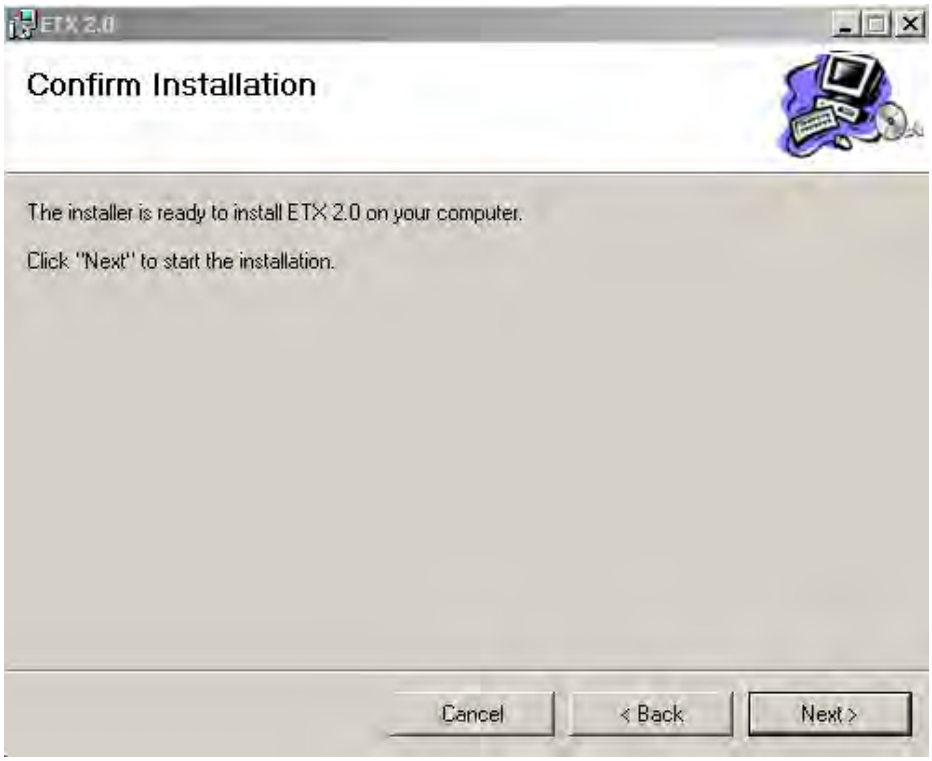


Figure 5. Installation procedure – Confirm Installation screen.

– In the dialog box displayed above, click **Next** if you want the installation to be carried out.
– Installation will now be performed. If the .NET framework is not yet installed on your computer, it will be installed. Installation of this framework takes several minutes.
– When the Installation is complete, the following screen appears. Click on **Close** to leave this screen and to finish the installation procedure.

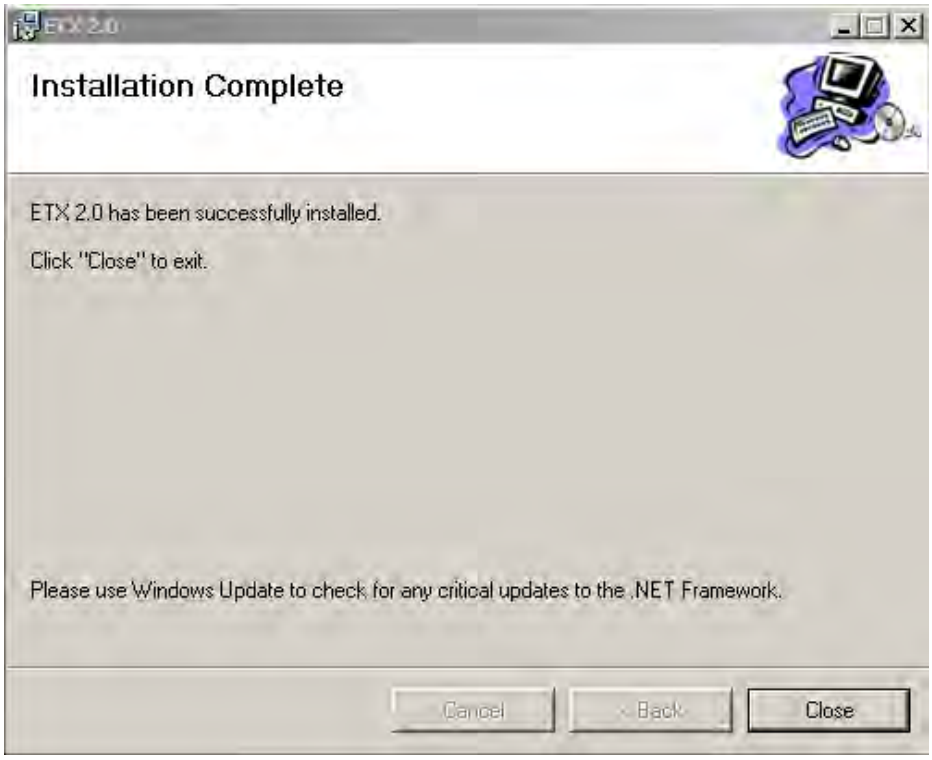


Figure 6. Installation procedure –Installation Complete screen.

3.3 Removing *ETX*

If you want to de-install (or remove) *ETX*, do the following.
- Under the Start button, click on Settings, Control Panel.
- In the Control Panel, double click the Add/Remove Programs icon.
- Select *ETX* from the list of currently installed programs. After clicking on *ETX* in this list, you can select Remove.
- *ETX* will now be removed from you computer.

3.4 Starting *ETX*

When *ETX* is installed successfully, the program can be found under the Start button in the lower left corner of your screen. It will be placed under: Programs, RIVM ETX, *ETX* 2.0. Click on the *ETX* 2.0 icon to start the program.

3.5 Saving data in *ETX*

If you save data in *ETX*, your data will be stored in a Project. This is a file with the extension: etx. To save a Project, Click on **File** on the Menu bar, then click **Save**.

In order to set or change the default directory where your *ETX* projects will be stored, see sections 5.1.5 and 5.9.4.

N.B. Your calculations and calculated results will *not* be stored together with your data. However, if you open an etx-Project that you had saved earlier, and press calculate again, all

results and output will be generated, together with the settings that were active at the time your Project was saved.

3.6 Starting a new project in *ETX*

The *ETX* opening screen is always an empty project. You can start directly with entering data here. Starting a new project can be performed by clicking on **File**, **New** in the Menu bar at the top of your *ETX* screen. Opening an existing project can be performed by clicking **File**, **Open input file**. The following dialog box appears:

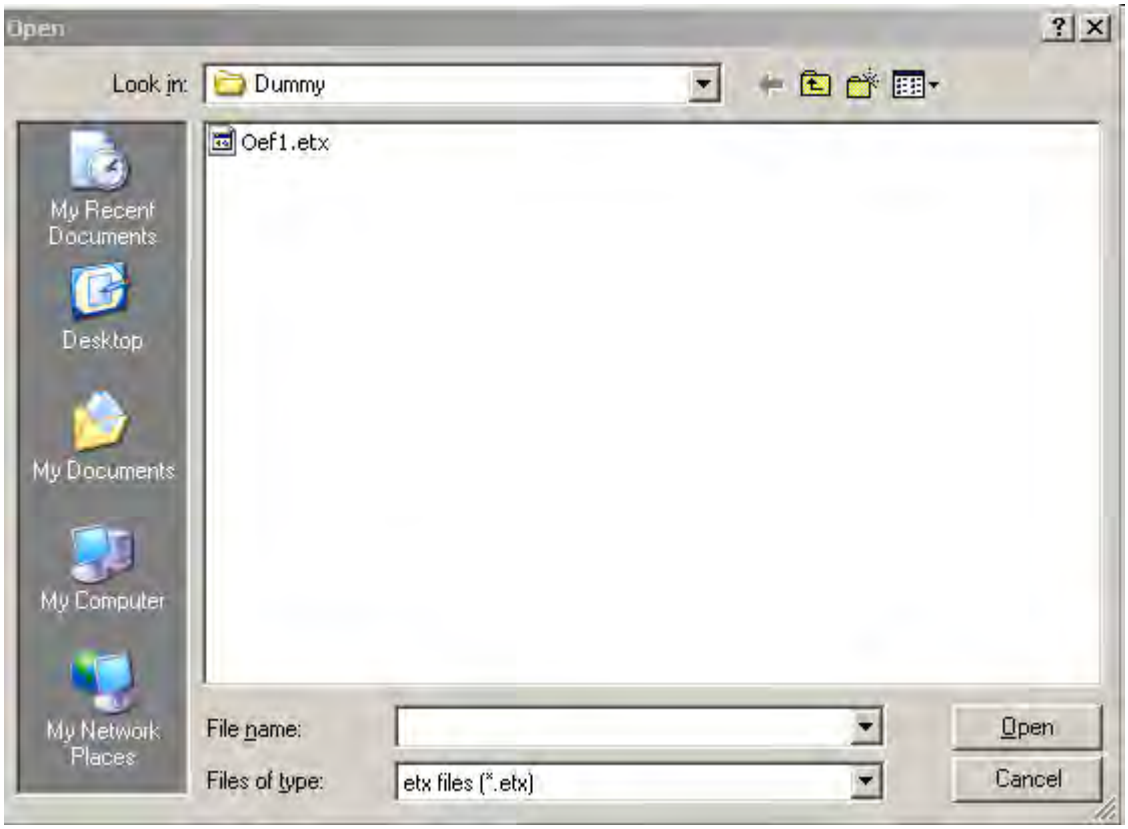


Figure 7. The dialog box shown after clicking **Open input file**.

Select the *ETX* project of you choice and click on **Open**. Your Project will now be retrieved and opened.

3.7 Leaving *ETX*

You can leave *ETX* 2.0 by choosing **File**, **Exit** from the Menu bar.

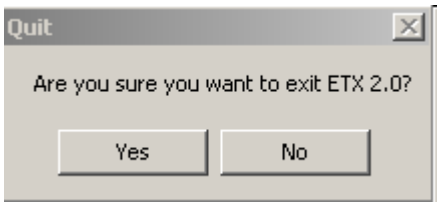


Figure 8. Dialog box shown upon leaving ETX.

Answer the question appearing in the dialog box (Figure 8) by clicking **Yes** or **No**.

3.8 Changing the settings for the decimal character

The choice of the character used as decimal symbol or digit grouping symbol in numerical values within *ETX* is determined by the regional settings of your PC. In English or American settings, the comma ',' is mostly used as digit grouping symbol, while the period '.' is used as decimal sign. In the Dutch language and settings, this is opposite, here the period '.' is used as digit grouping symbol and the comma ',' is the decimal sign. Please take notice that the regional settings will also apply to values printed along the axes of figures.

You may want to change these settings. Below, you find the route to the Control Box in which you can change your settings for several Microsoft operating systems.

3.8.1 Windows 98, NT, 2000 and ME

Click on the Start button in the lower left corner of your screen, then select Settings, Control Panel and Regional Settings (double click). In the Regional Settings properties box that has appeared, select the Number tab. In this screen you can either select a Region and obtain its accompanying settings or Customise your settings. After having made your selection, click on Apply and on OK. Changing of settings will have no effect unless you restart *ETX*.

3.8.2 Windows XP

Click on the Start button in the lower left corner of your screen, then select Settings, Control Panel, Regional and Language Options. Select the Regional Options tab. In this screen you can either select a Region and obtain its accompanying settings or Customise your settings. After having made your selection, click on Apply and on OK. Changing of settings will have no effect unless you restart *ETX*.

3.9 Comments on *ETX*

If you have comments on the program, its functioning or other ideas, you can send these by e-mail to etx.info@rivm.nl. Please note that this is *not* a helpdesk address. We will not respond to questions, but we will collect comments, possible errors and ideas for future development. Unfortunately, we are unable to help all users with problems or questions.

4. Reference manual

In section 4.1 we outline the different sections in which the program can be divided. We also name all sheets and briefly explain their contents. Section 4.2 explains the meaning of each of the items you may encounter in *ETX*. A list of abbreviations is included at page 63 of this report. How to use the information given in this chapter is explained to you in chapter 5 while chapter 7 tells you how to perform some calculations in practice.

4.1 Structure of the program

The program is divided in two major sections: an input and an output section. The input section has two subsections: **Input toxicity data** and **Input exposure data**. Each will be briefly discussed in the section 4.1.1. To get results, you have to invoke a calculation (section 5.7), which gives results in the **Output** section. The output section is divided in three sections, which will be introduced in section 4.1.3.

4.1.1 Data input

We discern two types of data: toxicity data and environmental concentrations.

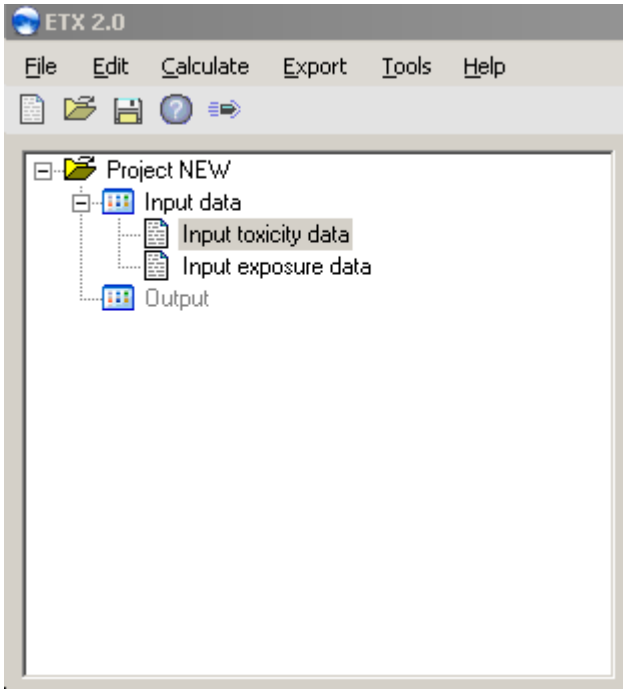


Figure 9. Subdivision of the input section.

4.1.1.1 Input toxicity data

Figure 9 shows the two sections of **Input data**.

- If you want to calculate an SSD and the accompanying HC_5 and HC_{50}, you should enter your data in the **Input toxicity data** sheet. If you want to calculate an FA, you need an SSD first. The output section will remain inaccessible until you have calculated an SSD.
- For calculation of the HD_5 for birds and mammals, there is room to enter a (small) data set of no more than 10 values. We call this the 'small sample' method. To use this method you have to select the check box labelled with **Use small sample method** in the lower right corner of the input screen (see Figure 10). You also have to enter a standard deviation

from an external data set. Entering data for the 'small sample' method is also done in the **Input toxicity data** sheet.

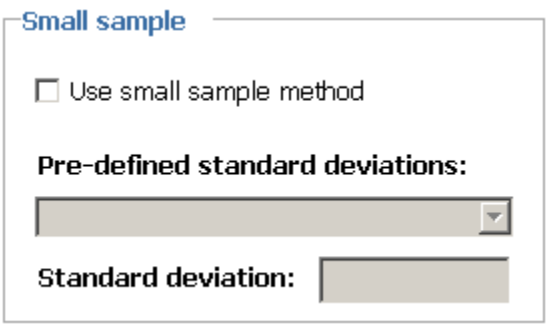


Figure 10. Small sample box.

4.1.1.2 Input exposure data

After having entered data in the **Input toxicity data** sheet and after you have calculated the accompanying SSD, you can use the **Input exposure data** sheet to calculate either an FA or an EER. There is the possibility for input of a single exposure concentration and a separate input section in case you have a series of exposure concentrations (like a measurement series).

4.1.2 Types of calculations

You can perform four types of calculations.

1. Calculate an SSD only.
2. Calculate an SSD and an FA. An FA can only be calculated when you have entered data in **Input toxicity data** and generated an SSD.
3. Calculate an SSD and an EER. An EER can only be calculated when you have entered data in **Input toxicity data** and generated an SSD.
4. Calculate a (small sample) HD_5. A maximum of 10 data can be entered. Calculation of an SSD is not needed in this case.

All calculations are performed by clicking **Calculate**, **Go** on the Menu bar in each of the two Input sections, or by clicking the **Calculate** button:  . Please note that pressing the **enter** key on your keyboard does *not* invoke a calculation.
After selecting the **Calculate** button (or using the menu), the calculation is performed and the results will appear in one or more of the output worksheets or graphs (see section 4.1.3).

NB. The program will also generate results and graphics if the outcome of the goodness-of-fit tests imply that it is less probable that your data derive from a normal distribution. The program only calculates, it does not make decisions. It is up to you to decide whether you accept or reject the presented outcome of the calculations.

4.1.3 Output

4.1.3.1 SSD related output

The **Output** section contains three subsections and each of the subsections contains one or more sheets. Figure 11 shows the available output sheets in the program tree below **Output**. The contents of each sheet are outlined below.

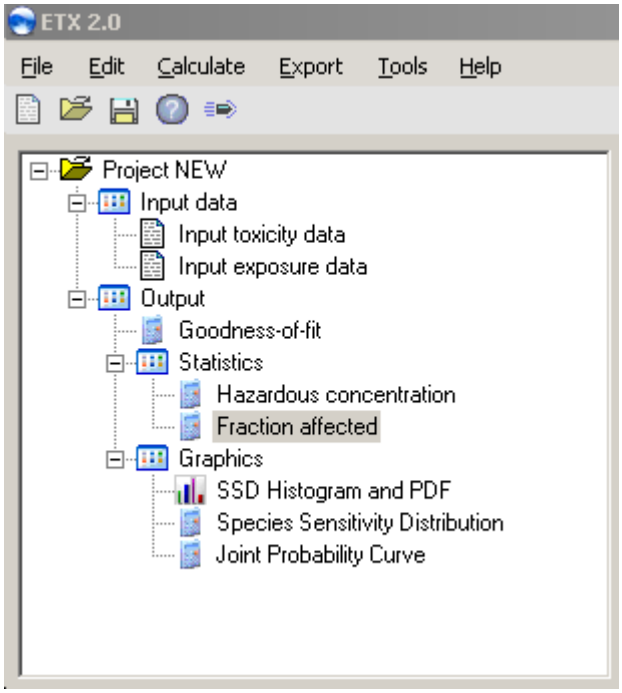


Figure 11. Subdivision of the output section for SSD calculations.

4.1.3.1.1 Goodness-of-fit

This sheet shows you the results of three goodness-of-fit (GOF) tests performed on your toxicity data. A brief explanation on the interpretation of these tests is also given. If you have entered a series of environmental concentrations in the **Input exposure data** sheet and have invoked a calculation, you will find the GOF test on these data here.

4.1.3.1.2 Statistics

In this section, two sheets can be found, the content of each is described below.

Hazardous concentration

In this sheet the mean and standard deviation (s.d.) of the normal distribution through your data are reported as well as the sample size. Furthermore, the HC_5, FA at the HC_5, the HC_{50} and the FA at the HC_{50} are reported. For each of these parameters, the lower and upper limit of the 90% confidence interval around the median estimate are reported.

Fraction affected

The estimated FA (plus lower and upper estimate of the two-sided 90% confidence interval) at the exposure concentration (EC) that you have entered in the **Input exposure data** sheet in the **single PEC** input cell is reported in this sheet. If you have entered a series of exposure concentrations in the **Input exposure data** sheet, the EER is reported.

4.1.3.1.3 Graphics

There are 3 sheets that show graphical output.

SSD Histogram and PDF

A histogram (or frequency distribution) of your toxicity distribution is presented when you have calculated an SSD. The bin width (classwidth) is calculated according to the method of Scott (1992) as:

$$binWidth = 3.5 \cdot \frac{standard\ deviation}{n^{1/3}}$$

The placement of bins on the *x*-axis is described in more detail in section 4.2.5. The height of the bars is expressed as the number of data they represent, which is plotted on the right hand *y*-axis. The normal distribution plotted in this graph is the probability density function (PDF) that is associated with your toxicity data sample. The left hand *y*-axis shows the density of the toxicity data. The left hand *y*-axis can also be used to read the densities of the bars in the histogram. For both the PDF and the histogram, the integral of all data, i.e. the sum of class-width × density, equals 1.

Species sensitivity distribution

If you have calculated an SSD, this is the cumulative density function (CDF) that is associated with your toxicity data sample. The dots in the graph are placed at the so-called Hazen plotting positions: $p_i = (i - 0.5)/n$ (see Aldenberg *et al.*, 2002).

Joint probability curve

This graph shows the joint probability curve (JPC). To obtain a JPC you have to enter toxicity data in the **Input toxicity data** sheet and a series (n>1) of exposure concentrations in the **Input exposure data** sheet. After having calculated an SSD, the JPC will be generated.

4.1.3.2 Small sample output

When the 'small sample' method is used, no SSD is required. The **Output** section only shows the **Small sample results** sheet (Figure 12). The content of the output sheet is outlined below.

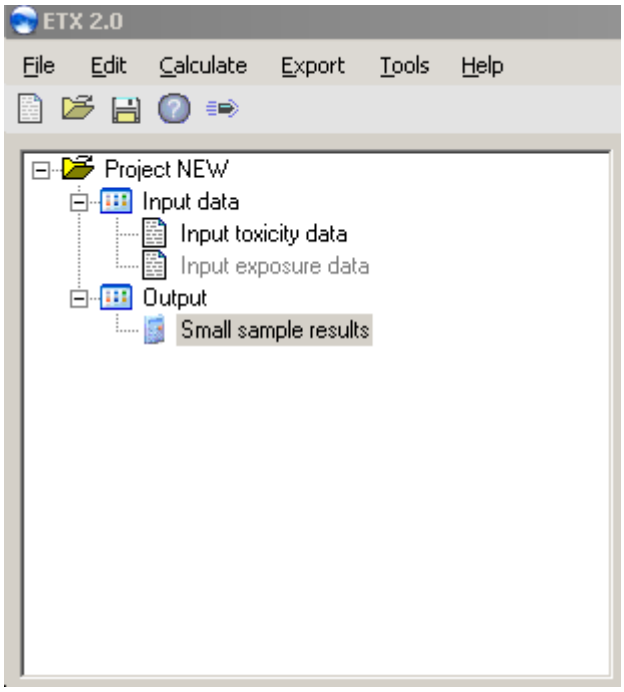


Figure 12. Subdivision of the output section for small sample calculations.

4.1.3.2.1 Small sample results

The median estimate of the HD_5 of your small data set is calculated and presented together with its 90% confidence interval limits. Also presented are the extrapolation factors that can be used to directly calculate the lower limit (LL), median or upper limit (UL) estimate of the HD_5 from the sample mean.

4.2 Glossary of keywords

In this section, a description of keywords is given, listed in alphabetical order. Keywords are those items you may encounter on the input or output screens of *ETX*. This section gives a brief explanation of each keyword, but we do not go deeply into the theoretical background of each item in this manual. For more background information, we refer to the underlying literature, which is listed in the Reference section of this manual.

4.2.1 Exposure concentration (EC)

In this manual and in *ETX* 2.0, the parameter 'exposure concentration' is synonym with 'environmental concentration'. Both can be abbreviated as EC and both are interchangeable in manual and program.

4.2.2 Expected Ecological Risk (EER)

An EER can be calculated if a set of predicted exposure concentrations (ECs) is entered in the input section of the program. Mean and s.d. are scaled relative to the mean and s.d. of the species sensitivity distribution. The EER literally is the probability that a randomly drawn species for a random draw of exposure is affected.

The following statistical parameters will appear in the **Fraction affected** output screen (in order of appearance):

SEC_{mean}, **SEC_{sd}** and **EER**

SEC stands for scaled environmental concentration.

- The SEC_{mean} is the scaled mean of the environmental concentration distribution. It is reported here as the $\log_{10}$ transformed value.
- SEC_{sd} is the scaled s.d. of the environmental concentration distribution. It is reported here as the $\log_{10}$ transformed value.
- EER is the expected (mean) ecological risk, given the SSD and the distribution of environmental concentrations. The EER is the area below the curve in the **joint probability curve** (**Output**, **Graphics**).

4.2.3 Extrapolation factor

In the **Small Sample Output** sheet, extrapolation factors are displayed. These can be found in the **HD_5 (median) results** box on your screen.

Extrapolation factors are factors to be applied multiplicatively to the geometric mean of the original (not log-transformed) toxicity data. Extrapolation factors depend on the FA, sample size, confidence level and on the standard deviation of the SSD (Aldenberg and Luttik, 2002). In *ETX* 2.0, extrapolation factors are given for an FA of 5% (HD_5) for the upper and lower limit of the 90% confidence interval of the HD_5.

4.2.4 Fraction affected (FA)

An FA can be calculated when one single predicted exposure concentration (EC, synonym with PEC) value is available. This EC value can be entered in the input section of the program. The following statistical parameters will appear in the **Output** screen. In order of appearance:

LL FA, Median FA and **UL FA**

Reported is the median estimate of the FA at the PEC you have entered. Also reported is the two-sided 90% confidence interval of the FA: LL FA represents the 5% confidence limit and

UL FA represents the 95% confidence limit. These calculations are based on repeated linear interpolation and are approximate values. The procedure is explained in section 8.2.3 of Aldenberg and Jaworska (2000). A precise answer can be obtained using the functions described in section 8.2.1 of that paper.

This routine does not yield results when your standardised mean logarithmic effect concentration is <-5 or >5. As a check, the value of the standardised logarithmic concentration calculated from your PEC value is shown in the **Fraction affected** output sheet. If this value is outside the range mentioned above, no results will be given. The message
'The method does not yield results because your standardised PEC exceeds the limits of <-5 or >5' will appear in red. The sheet will show 'Out of bounds!' as error values.

4.2.5 SSD Histogram and PDF

This graph is a histogram or frequency distribution in which the toxicity data are represented by bars. On the *x*-axis the $\log_{10}$ of toxicity values (NOEC, LC_{50}s etc.) is plotted. On the right hand *y*-axis, the frequency of the toxicity values is plotted. Toxicity values are distributed over frequency classes (often referred to as bins) that have a width calculated according to Scott (1992):

$$binWidth = 3.5 \cdot \frac{standard\ deviation}{n^{1/3}}$$

The placing of bins on the *x*-axis is worked out as follows:
1. The number of bins is calculated using the equation mentioned above;
2. Bins are divided over the range of toxicity data by centering around the mean (=sample mean, or mean of the toxicity data);
3. Sample values coinciding with a bin limit are dropped in the next lower bin.
4. In case the lowest sample value coincides with the lower limit of the lowest bin, this value is dropped in this lower bin.

In case of an even number of bins, we identify *two* middle bins. The mean of the toxicity data now coincides with the separation between the two middle bins. More precise, the mean value itself is chosen to coincide with the highest value of the left bin (due to point 3 above). In case of an uneven number of bins, there is *one* middle bin. In these cases, the middle of the middle bin is chosen to coincide with the mean of the toxicity data.

4.2.6 Goodness-of-fit

Whether the sample of toxicity data derives from a normal distribution can be assessed with goodness-of-fit tests. Two different types of tests are implemented, based on quadratic (vertical) distance and on the largest vertical distance. Well-known quadratic tests are the Anderson Darling test and the Cramér-von Mises test (cf. D'Agostino and Stephens, 1986). The Kolmogorov-Smirnov test is a well-known vertical distance test (D'Agostino and Stephens, 1986).

1. The Anderson-Darling goodness-of-fit test highlights differences between the tail of the distribution and the input data and is generally regarded as a very powerful general test (Aldenberg *et al.*, 2002).
2. The Kolmogorov-Smirnov test focuses on differences in the middle of the distribution and is not very sensitive to discrepancies of fit in the tail of the distribution.

Interpreting Critical Values

If a test statistic is above the 5% critical value, normality is rejected at the 5% critical value, indicating doubts about normality.

If a test statistic is below the 5% critical value, normality is accepted (not rejected) at the 5% critical value.
If a higher critical value is accepted (e.g. at 2.5% significance level), then the probability that these data derive from a normal distribution is smaller than at 5%, but it is not impossible that the sample derives from a normal distribution.
Some people find it confusing that a higher significance level, say 10%, has a lower critical value. 'Rejected' occurs more often with a higher significance level. In conclusion: a GOF test does NOT say that a sample cannot derive from a normal distribution, just that it becomes less probable with decreasing significance levels.

4.2.7 HC_5 and HC_{50}

Your toxicity dataset refers to a selection of species, which is treated as a sample drawn from a population (in a statistical sense). By definition, the calculated normal distribution encompasses all species inhabiting the environmental compartment of interest. It is up to you to determine whether your sample of species is representative for the population of species (in a biological sense) you want to derive HC values for. The distribution is the –estimated– function that relates the relative sensitivity of the species thought present in a given environmental compartment to (the logarithm of) the toxicant concentration. This distribution is a normal distribution (by definition) when you use *ETX* 2.0 for your calculations. HC is short for 'hazardous concentration'. HC_5 and HC_{50} are the 5^{th} and 50^{th} percentile (median is synonymous for the latter) of the normal distribution that is fitted through the toxicity data you have entered. HC_5 and HC_{50} are expressed as a toxicant concentration (in the same units as the toxicity data you have entered in the input section). Hence, they represent the toxicant concentration that is hazardous to 5 or 50 percent of 'all' species. See section 4.2.11 for explanation of the HC parameters calculated by *ETX* 2.0.

4.2.8 JPC

The abbreviation stands for joint probability curve. This is a graphical representation of the risk of a substance to the species, that may be used in risk characterisation. The type of graph *ETX* shows is a cumulative profile plot. It is constructed by plotting FA values from the SSD (in CDF form) on the *y*- axis against exposure concentration distribution (ECD) values (also in CDF form) on the *x*-axis at corresponding log-exposure concentrations. The area under the curve (AUC) is equal to the expected ecological risk. For more detail, see Aldenberg *et al.* (2002). The numerical value of the AUC, or EER, is shown in the **Fraction affected** sheet (**Output**, **Statistics**).

4.2.9 PEC

If an environmental concentration of the substance of interest is the outcome of some model calculation rather than a measured value –or series of values– in the field, it is called a *predicted* environmental concentration. If you have a single PEC value and an SSD, you can calculate the fraction of species described by that SSD that is (potentially) affected by the substance.

4.2.10 Small Sample method

The idea to estimate a percentile of an SSD and its uncertainty for a very small sample of toxicity data has emerged from the field of pesticide registration. Due to the limited size of the sample, there is no reliable information on the standard deviation of the SSD. A standard deviation from a *different* sample, to which one attaches more confidence because it has higher reliability, is used to estimate the HD_5. Luttik and Aldenberg (1997) published this method for *logistically* distributed toxicity data.

The method that is currently implemented in *ETX* 2.0 offers the possibility to calculate HD_5 values for birds or mammals based on *normally* distributed toxicity data (Aldenberg and Luttik, 2002). From a dataset of 55 pesticide LD_{50} values for birds a pooled variance estimate of the standard deviations of all LD_{50} values was calculated. The same was done for a dataset of 69 pesticide LD_{50} values for mammals. From both datasets the pooled variance estimates for carbamates and organophosphorous compounds were also calculated. All pooled variance estimates are listed in the **Input toxicity data** sheet in the **Small sample** box (under **Pre-defined standard deviations**). These estimates can be used as an ('external') standard deviation that is assumed to describe the spread of the population from which your (small) sample of bird/mammal data has been drawn.

Recommendations to the 'small sample' method:
(i) When there are indications that the (small) sample standard deviation does not reflect the population standard deviation, even when $n \geq 4$, consider using the pooled standard deviation (i.e. use the 'small sample' technique). A maximum of 10 entries is allowed for in the **Input toxicity data** sheet.
(ii) Use the LL HD_5 values or corresponding assessment factors when a 5% probability for overestimation of the HD_5 is desired. When the median HD_5 or its corresponding assessment factor is used, 50% probability for HD_5 overestimation is allowed.
(iii) When there are indications that the available data are derived from a test with a sensitive species, one could consider using the median estimate of the HD_5 or the corresponding extrapolation factor.

4.2.11 Statistics SSD

If an SSD has been calculated by *ETX*, i.e. if a normal distribution has been fitted through your toxicity data, the following statistical parameters will appear in the **Hazardous concentration** output screen (in order of appearance).

Mean, s.d. and ***n***

Mean and s.d. are the sample mean and the sample standard deviation (n-1) of the normal distribution. These parameters are shown in $\log_{10}$ units and are the parameters that describe the normal distribution fitted through your toxicity data set. n is the sample size, i.e the number of toxicity data. It is shown because the size of n directly determines the size of the uncertainty in the HC_5 and HC_{50} (Aldenberg and Jaworska, 2000). These sample statistics are used to estimate the SSD parameters.

HC_5 results: **LL HC_5, HC_5, UL HC_5** and **sprHC$_5$**

These parameters are shown in the same concentration units as you entered your toxicity data in. Presented are the estimated 5th percentile of the normal distribution (HC_5) and its two-sided 90% confidence limits, called the lower limit (LL) and upper limit (UL), respectively. The HC_5 itself is a normally distributed statistic, and sprHC$_5$ is a measure of the width of the HC_5 distribution. It is calculated as the ratio of UL and LL.

FA at HC_5 results: **FA_{lower}, FA_{median}** and **FA_{upper}**

The median estimate of the FA at the toxicant concentration HC_5, is 5%, by definition, since the HC_5 is the 5th percentile of the SSD. The two-sided 90% confidence interval for the FA is also reported: FA_{lower} represents the 5% confidence limit of the FA and FA_{upper} represents the 95% confidence limit.

HC_{50} results: **LL HC_{50}, HC_{50}, UL HC_{50}** and **sprHC$_{50}$**

These parameters are shown in the same concentration units as you entered your toxicity data in. Presented are the estimated median or 50th percentile of the normal distribution (HC_{50}) and its two-sided 90% confidence limits, called the lower limit (LL) and upper limit (UL), respectively. The HC_{50} itself is a normally distributed statistic, and $sprHC_{50}$ is a measure of the width of the HC_{50} distribution. It is calculated as the ratio of UL and LL.

FA at HC_{50} results: **FA_{lower}, FA_{median}** and **FA_{upper}**
The median estimate of the FA at the toxicant concentration HC_{50}, is 50%, by definition, since the HC_{50} is the 50th percentile of the SSD. The two-sided 90% confidence interval for the FA is also reported: FA_{lower} represents the 5% confidence limit of the FA and FA_{upper} represents the 95% confidence limit.

4.2.12 SSD

E_TX calculates an SSD from your sample of toxicity data assuming that the (continuous) distribution describing the sensitivity of the population underlying your sample is normally distributed over log concentration. This Gaussian distribution is graphically presented in the **SSD Histogram and PDF** sheet (to be found under **Output**, **Graphics**). The data and the fitted distribution are also presented as a CDF plot in the **SSD** sheet (**Output**, **Graphics**). The statistical parameters derived from this distribution are presented in the **Hazardous concentration** sheet (**Output**, **Statistics)**.

5. Working with *ETX*

5.1 Menu bar and buttons

The menu bar and buttons that are available in *ETX* are shown in Figure 13. Throughout the program, the same menu bar and buttons will appear. All functions that may be accessed via the menu bar will not be described here individually, we will give a general description of each menu item. Most items you will encounter when you work through the manual or by just trying them!

Figure 13. Menu bar and buttons in ETX.

5.1.1 File

Options provided here allow you to:

- start a new *ETX* project (**New**).
- open an existing *ETX* project (**Open input file**). When you have saved your work in an *ETX* project, *ETX* does *not* save the calculated data and graphs. If you want to see the results again, simply press Calculate after opening an existing *ETX* project and your results will reappear. If you want to save all data, choose Export and all your results will be placed in an Excel spreadsheet.
- save your work (**Save**).
- save your project under a different name (**Save as**).
- leave *ETX* (**Exit**).

5.1.2 Edit

The options provided here allow you to handle individual cells or ranges of cells, that you can either delete, cut, copy and paste.

5.1.3 Calculate

The only option provided here is to invoke a calculation. The function of this menu is equal to that of the calculate button (section 5.1.7). For more information on the types of calculations you can perform, see section 5.7.

5.1.4 Export

The only option provided here is to export the data you have entered and calculated, to a separate file. The data will be exported in Microsoft Excel spreadsheet format. See also section 5.8. To change the default directory where projects and exported files are stored, see section 5.9.

5.1.5 Tools

There are four options here: **Labels**, **Sort toxicity data**, **Sort labels** and **Preferences**.

- Use **Labels** to attach label to the entries in the toxicity data sheet. For more information please see section 5.6.
- By clicking on **Sort toxicity data,** the toxicity data in the **Input toxicity sheet** will be sorted in increasing order. NB: this operation can not be undone.

- By clicking on **Sort labels,** the labels in the **Input toxicity sheet** will be sorted in alphabetical order. NB: this operation can not be undone.
- Use **Preferences** to define the default settings for the looks of your graphs and to set the default directory where projects and exported files are stored. For more information please see section 5.9.

5.1.6 Help

In the **Help** menu, there are two options to choose: ***ETX* Help topics** and **About *ETX* 2.0.** When you select the help topics, you can use the *ETX* 2.0 manual online. The content of this menu is equal to the content of this manual.

About *ETX* 2.0 gives you information on the version of this program. Information on earlier versions of *ETX* is also provided.

5.1.7 Buttons

The buttons have the following names and function:


New. This button opens a New *ETX* project.


Open existing project.


Save. Allows you to save the current project.


Help. Allows you to access the Help file.


Calculate. Invokes a calculation.

5.2 Navigating through *ETX*

After starting up *ETX*, the following screen will appear:

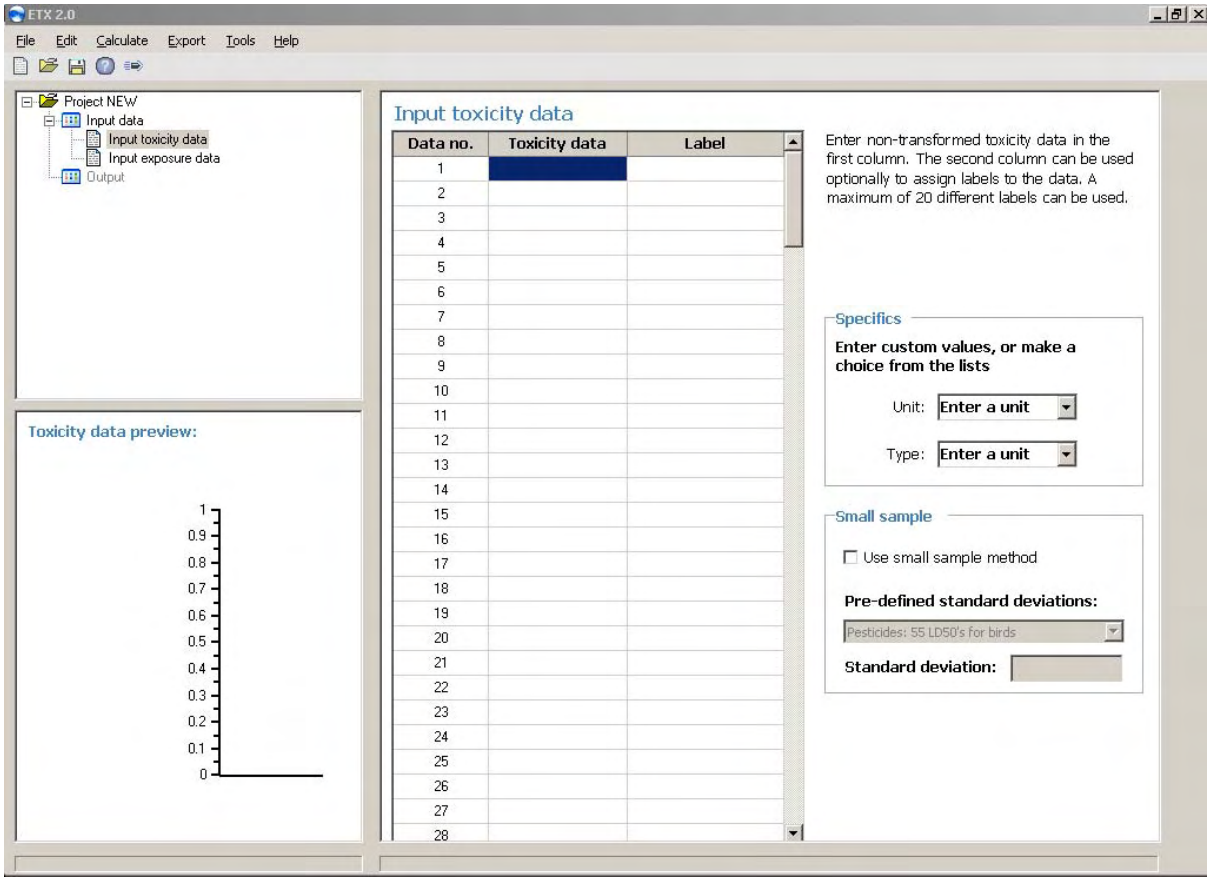


Figure 14. E_TX appearance after starting the program.

This screen is divided in three major parts: in the upper left corner: the navigating section, in the lower left corner: the **Toxicity data preview** section and the largest part is the right half of the screen: the **Input toxicity data** section. In the navigating screen, you find + or – signs. By clicking on these you can *unfold* or *fold* a specific part of the program and *view* or *hide* its contents. Use these buttons to find out that each E_TX project contains an **Input data** section and an **Output** section. If you select a specific section or sheet by clicking on it, it will be displayed in the right half of your screen. In the following, you will be guided through each of the different parts of the program.

5.3 Entering data

5.3.1 Toxicity data

- Go to the **Input toxicity data** sheet in the **Input data** section of E_TX.
- Go to the **Input toxicity data** box (this is the right part of your screen).
- Next, in the cells under **Toxicity data**, enter your toxicity values as a column, with one value per input cell. Enter the toxicity data in the original (non-transformed) values.
- For information on entering the unit of your data or the type of endpoint, see section 5.3.2.
- For information on labelling of your toxicity data, see section 5.6.
- Make sure *all* entries have the same unit: mg/l, µg/kg, ppm etc.
- You cannot enter the value 0 (zero). Your data will be $\log_{10}$-transformed for calculation and log(0) does not exist. If you enter a zero value, the error message **Please enter only**

positive values will be displayed at the bottom of your screen. Delete the zero value or enter a positive non-zero value to continue.
- You need not sort the data.
- You may leave blank cells, empty cells will not be included in the graph or in calculations.
- For HC_5 and HC_{50} calculation a maximum number of 200 data can be entered.
- For FA and FA-confidence limits the maximum number of data (n) =75 for FA at HC_{50} and n = 200 for FA at HC_5.

When you have finished entering data, see section 5.7.1 on how to calculate an SSD.

5.3.2 Assigning units to your data

You may enter the unit of toxicity data and/or exposure data. It is optional to do this, but there are two reasons to do so. First, if you are about to enter both toxicity and exposure data e.g. to calculate a fraction affected, the data should be entered in E_TX in the same units! Second, if you enter the unit of your data, this unit will be printed in the output of your data that is exported to Excel.

To assign a unit to your data go to the **Input toxicity data** section. In the lower right half of your screen, you find the **Specifics** box (Figure 15):

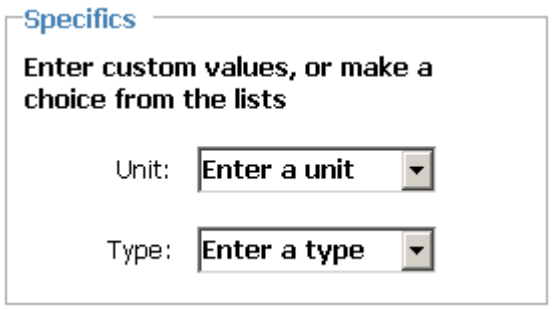


Figure 15. The specifics box.

- In the scrollbar to the right of **Unit**, you may type a unit of your choice or select the unit from the pulldown menu.
- In the scrollbar to the right of **Type**, you may type the endpoint of the toxicity data you are about to enter, or you can select an endpoint from the pulldown menu.
- The **Specifics** box functions as a piece of scrap paper. Both options in the box remind you that all toxicity data should be of the same unit and type. Unit and type are not linked to any calculation. They reappear in the **Input exposure data** section, to remind you that the units of data entered in that section should be identical to the units of the data entered in the **Input toxicity data** section.
- Use of the Specifics box is optional. E_TX works equally well when you clear the contents of the **Unit** and **Type** boxes.

5.3.3 Small sample data

- Go to the **Input toxicity data** sheet in the **Input data** section of E_TX.
- In the **Input toxicity data** box (this is the right part of your screen), click on the **Use small sample approach** check box in the **Small sample** box. The following dialog box will appear:

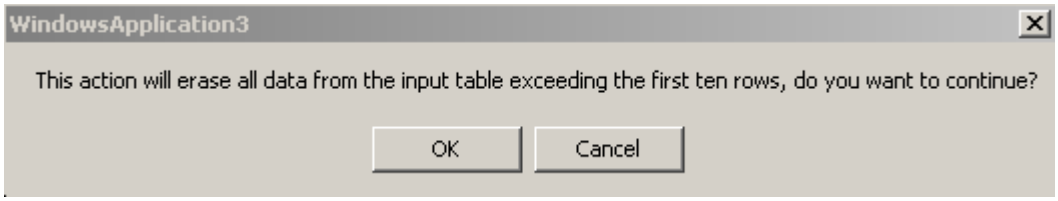


Figure 16. Small sample verification box.

If you had already typed data in the input section, these data will be lost when you switch to the Small sample input mode. Remember that this is a different method that does not require calculation of an SSD. So, if you want to keep the data you had entered before, click on **Cancel** and Save your work in an *ETX* project. Then start a new project for your 'small sample' calculation. Otherwise, click on **OK**.

- The number of toxicity data that you can enter is now limited to 10.
- Enter your toxicity values as a column, with one value per input cell. Enter the toxicity data in the original (non-log transformed) format.
- Make sure *all* entries have the unit of mg/kg body weight.
- You cannot enter the value 0 (zero). Your data will be $\log_{10}$-transformed for calculation and log(0) does not exist.
- You need not sort the data.
- You may leave blank cells, empty cells will not be included in calculations.
- Enter a standard deviation. In the **Small sample** box at the lower right corner of your input screen, you can choose one of several pre-defined pooled standard deviations. The accompanying value that will be used for calculations appears in the **Standard deviation** box.

When you have finished entering data, see section 5.7.2 on how to calculate an HD_5.

5.3.4 Environmental concentrations

5.3.4.1 One environmental concentration

This calculation applies when you have only one environmental concentration, e.g. a PEC resulting from the exposure part of a risk assessment rather than a series of measurements.

- Go to the **Input exposure data** sheet in the **Input data** part of *ETX*.
- You need to have an SSD before you can enter exposure data. Refer to section 5.3.1 for information on this subject. If you enter exposure data without having an SSD, the following message will pop up:

Figure 17. Message displayed upon FA calculation when no toxicity data have been entered.

- Enter your PEC value in the **Single PEC** box at the right part of your screen.
- Make sure that the unit of your PEC value is identical to the unit of the toxicity values that you have entered to generate your SSD! The unit of your toxicity data will be displayed in the **Specifics** box when you have made use of it while entering of your toxicity data.

- Enter a non-log transformed PEC value!
- This routine does not yield results when your standardised logarithmic PEC is <-5 or >5. The value of the standardised logarithmic concentration PEC will be shown in the **Fraction affected** sheet (**Output**, **Statistics**).
- You cannot enter the value 0 (zero). Your data will be $\log_{10}$-transformed for calculation and log(0) does not exist.

When you have finished entering data, see section 5.7.3 on how to calculate an FA.

5.3.4.2 A series of environmental concentrations

This calculation applies when you have a series of environmental concentrations, e.g. a measurement series of the compound of interest in an environmental compartment.

- Go to the **Input exposure data** sheet in the **Input data** part of *ETX*.
- Go to the **Input environmental concentrations** box (this is the right part of your screen).
- In the cells under **Exposure data**, enter your toxicity values as a column, with one value per input cell.
- Enter the toxicity data in the original (non-transformed) values.
- Make sure that the unit of your environmental concentrations is identical to the unit of the toxicity values that you have entered to generate your SSD! The unit of your toxicity data will be displayed in the **Specifics** box when you have made use of it during entering of your toxicity data.
- Make sure *all* entries have the same unit: mg/l, µg/kg, ppm etc.
- You cannot enter the value 0 (zero). Your data will be $\log_{10}$-transformed for calculation and log(0) does not exist.
- You need not sort the data.
- You may leave blank cells, empty cells will not be included in the graph or in calculations.

When you have finished entering data, see section 5.7.4 on how to calculate an EER and JPC.

5.4 Clearing the contents of the input cells

5.4.1 Clearing a single cell

You can clear (or delete) data in a single input cell as follows.

- Select the cell you want to by clicking in the cell.
- Click on **Edit**, **Delete** in the Standard menu bar or press the Delete button on your keyboard to clear the content of the selected cell. The following message will appear:

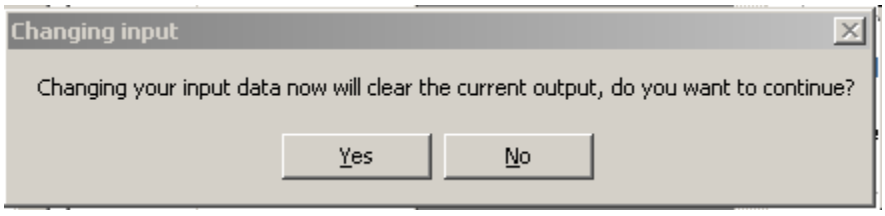


Figure 18. Message displayed when cell contents are about to be deleted.

- Click on **Yes** if to continue. The content of the cell will be deleted.

5.4.2 Clearing all cells

In order to clear a range of cells, do the following:
- Select the uppermost cell of the range of cells you want to clear, by clicking in the cell.
- Press the Shift button together with the Arrow Down (↓) button (both on your keyboard). While holding the shift button pressed down, you can select multiple cells by using the ↓ button repeatedly.
- When you have reached the bottom cell of the range you want to clear, stop holding down keys. Your selection of cells is now highlighted.
- Click on **Edit**, **Delete** in the Standard menu bar or press the Delete button on your keyboard to clear the content of the selected cells.

In order to clear all cells, do the following:
- Select the upper cell by clicking in it.
- Press the Shift button together with the Ctrl button. While holding these buttons pressed down, click on the Arrow Down (↓) button (all buttons are on your keyboard).
- All cells will now be selected (highlighted).
- Click on **Edit**, **Delete** in the Standard menu bar or press the Delete button on your keyboard to clear the content of the selected cells.

5.5 Importing data

You can also enter data (both toxicity data and environmental concentrations) from e.g. MS Excel. The procedure is as follows:
- The data you want to import should be placed in a column.
- Select the column containing your data and copy it
- Switch to *ETX* and click once in the upper cell of the data column of the sheet in which you want to import your data.
- Paste the data.

Please refer to section 5.11 on how to use the decimal symbol in these type of operations.

5.6 Adding labels to toxicity data

NB. This function is optional. All other program functions can be executed equally well when the toxicity data are not labelled.

When entering toxicity data, you have the possibility to label your entries. This means that you can add a text label to toxicity data. You might, for example, want to discriminate between the different taxonomic levels in your SSD and add labels like *algae*, *crustacea*, *insecta* (or abbreviations). Apart from adding text labels, you can customise the appearance of the symbols that are linked to your labels. These symbols will be used to construct your SSD.

5.6.1 Creating a label collection and using labels

Before you can enter labels that will also appear in your SSD graph, you have to create a label collection. To create a label collection, do as follows:
- Click on **Tools, Labels** in the Menu bar and the **Input toxicity data labels** dialog box shown in Figure 19 will appear:

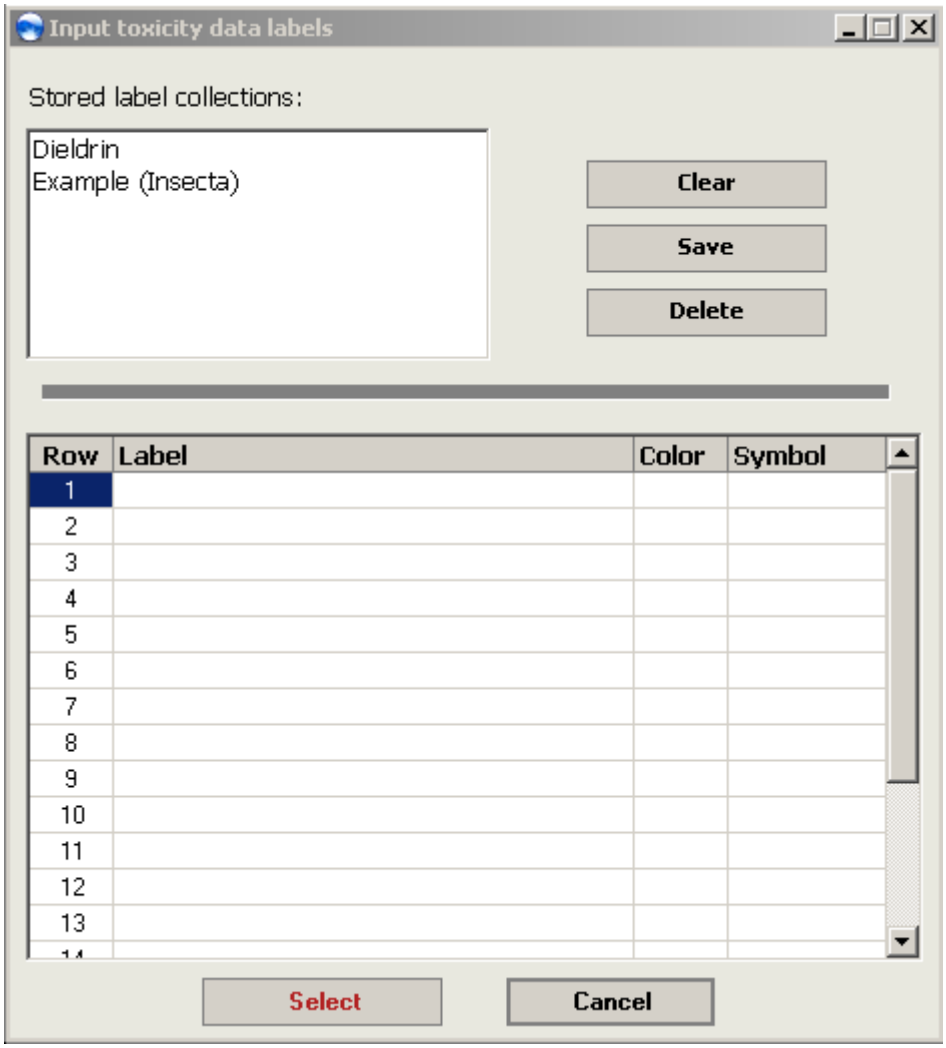


*Figure 19. The **Input toxicity data labels** box.*

- Next, type the name of the label in the column **Label** and assign a **Color** and a **Symbol** to the label in the following columns. Complete your label collection in this way.
- After having completed your label collection, click the **Save** button in the top part of the **Input toxicity data labels** box and type the name of your label collection. The name will now appear in the **Stored label collections** box visible at the top of the dialog box.
- You can select a label collection by clicking on its name in the **Stored label collections** box and subsequently press the **Select** button at the bottom of the label box. You then automatically return to the **Input toxicity data** sheet.
- Back in the **Input toxicity data** screen you see the selected label collection displayed in the lower right corner of your screen:

Figure 20. The active label collection display.

- Back in the **Input toxicity data** screen, behind each of the toxicity data that you enter, there is a **Label** cell. Using the pull down menu in each cell, you can now choose one of the

taxonomic groups (or 'labels') that are present in the label collection you have just selected.

5.6.2 Editing an existing label collection

Editing an existing label collection can also be performed via the **Input toxicity data labels** dialog box:

- Click on **Tools**, **Labels** in the menu bar and the **Input toxicity data labels** dialog box shown in Figure 19 will appear.
- Select the label collection in the **Stored label collections** box by clicking on it.
- Edit the label collection.
- Press the **Save** button to save any changes.
- Press the **Select** button if you wish to select the current label collection. You will automatically leave this dialog box after clicking **Select**.

5.6.3 Deselecting a label collection

In case you want to clear the connection between the toxicity data you have entered and the selected label collection, do the following:

- In the **Input toxicity data** screen you see the **Label collection display** in the lower right corner of your screen (Figure 20).
- Click on the **Deactivate** button to clear all labels entries behind your toxicity data.
- Note that this action will not erase your label collection or its settings, only its connection with the current dataset.

5.7 Performing calculations

5.7.1 Calculating an SSD

When you have finished entering toxicity data in the **Input toxicity data** sheet, you are ready to calculate an SSD.

- Select **Calculate**, **Go** from the Standard menu bar or press the **Calculate** button: to invoke the calculation of the SSD and its statistics.
- After performing the calculation, *ETX* will automatically switch to the **Hazardous concentration sheet** in the **Output, Statistics** section.
- Switch manually to the **Goodness-of-fit** section in order to check the result of the normality tests performed on your toxicity data.
- Two graphs will be generated as well: the histogram and an SSD. Both can be found under **Output, Graphics**.
- If you have deleted or added data in the input section, choose **Calculate**, **Go** from the Standard menu bar (or press) again for an update of statistics and graphics.

5.7.2 Calculating a small sample HD_5

When you have finished entering toxicity data in the **Input toxicity data** sheet using the 'small sample' method, you are ready to calculate an HD_5.

- Select **Calculate**, **Go** from the Standard menu bar or press the **Calculate** button: to invoke the calculation of the HD_5 and its statistics.
- After performing the calculation, *ETX* will automatically switch to the **Small sample results sheet** in the **Output, Statistics** section.

– If you have deleted or added data in the input section, choose **Calculate**, **Go** from the Standard menu bar (or press ) again for an update of statistics.

5.7.3 Calculating an FA

When you have calculated an SSD and finished entering exposure data in the **Input exposure data** sheet, you are ready to calculate an FA.

– Select **Calculate**, **Go** from the Standard menu bar or press the **Calculate** button:  to invoke the calculation of the FA and its statistics.
– After performing the calculation, *ETX* will automatically switch to the **Fraction affected sheet** in the **Output, Statistics** section.
– If you have deleted or changed your PEC in the input section, choose **Calculate**, **Go** from the Standard menu bar (or press ) again for an update of statistics.

5.7.4 Calculating an EER and JPC

When you have finished entering a series of environmental concentrations in the **Input exposure data** sheet, you are ready to calculate an EER and JPC.

– Select **Calculate**, **Go** from the Standard menu bar or press the **Calculate** button:  to invoke the calculation of the EER and its JPC.
– After performing the calculation, *ETX* will automatically switch to the **Fraction affected sheet** in the **Output, Statistics** section. This sheet shows the EER in the lower one of the two boxes.
– Switch manually to the **Goodness-of-fit** section in order to check the result of the normality test performed on your exposure data.
– A joint probability curve will be generated as well. It can be found under **Output, Graphics**.
– If you have deleted or added data in the input section, choose **Calculate**, **Go** from the Standard menu bar (or press ) again for an update of statistics and the JPC.

5.8 Exporting results

Once you have generated results and statistics, you might want to save these results or store them in some other format. This can be done by using the **Export** function. You can find the **Export** function in the menu bar. Clicking on Export shows you one menu option: **Export current output**. After selecting this option, your data, statistical output en graphs will be exported to an MS Excel spreadsheet (extension .xls). The default filename is identical to the name of your *ETX* project. If you have not yet stored you data in an *ETX* project, the default filename will be NEW.xls. The following dialog box appears:

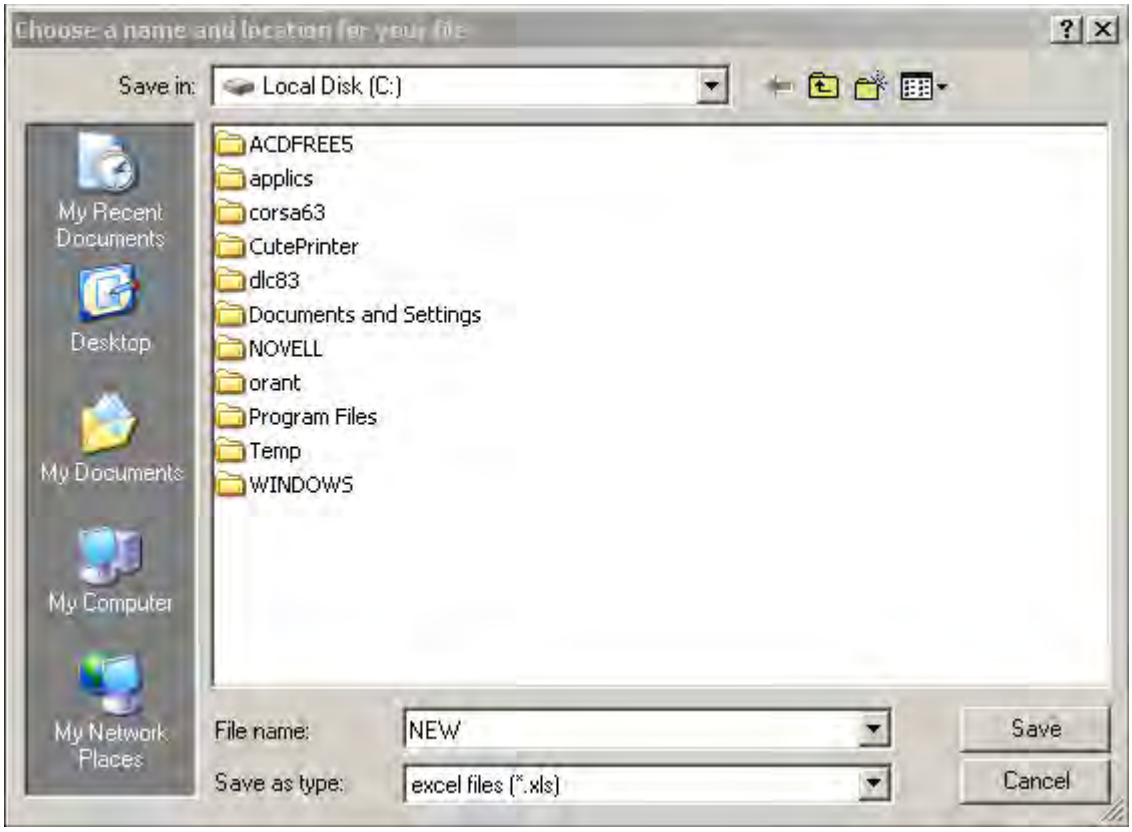


Figure 21. Save export file dialog box.

After having selected the location of your choice, click on **Save**. You will be asked if you want to open the export file directly or not:

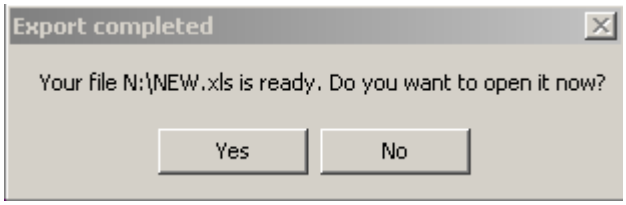


Figure 22. Export completed dialog box.

In order to have the exported (Excel)filenames corresponding with your E_TX-project names, save your E_TX data in an E_TX-project before exporting data to an Excel sheet. The name of your E_TX-project will then be chosen as default filename for your exported data file.

5.9 Preferences

5.9.1 Changing graph appearance in the *current* project

The three graphs that may be generated with E_TX 2.0 can be edited to a limited extent. This section explains how.

- Go to one of the graphs that you have generated: **SSD Histogram and PDF**, **Species sensitivity distribution** or **Joint probability curve** under **Graphics**.
- In the graph of your choice, click with the right mouse button (left button for left handed mousers).
- A pop up-window appears in which you can select various parts of the graph that you might want to edit (e.g. Title, Title font, labels, line colour etc.).

– All changes that you have made using this option will be saved along with your project. When you re-open a project and press calculate, the settings you last made will appear in your graphs.

5.9.2 Changing the axis settings

There are three possible graphs in E_TX: an **SSD Histogram and PDF**, a **Species sensitivity distribution** and a **Joint probability curve**. The axis settings of the **SSD Histogram and PDF** can not be changed. Axis settings for the other two graphs can be changed as follows:

– Go to one of the graphs that you have generated: **SSD Histogram and PDF**, **Species sensitivity distribution** or **Joint probability curve** under **Graphics**.
– In the graph of your choice, click with the right mouse button (left button for left handed mousers).
– A pop up-window appears in which you can select various parts of the graph that you might want to edit (e.g. Title, Title font, labels, line colour etc.). Select **Axes settings** from the appearing menu. The following dialog box appears:

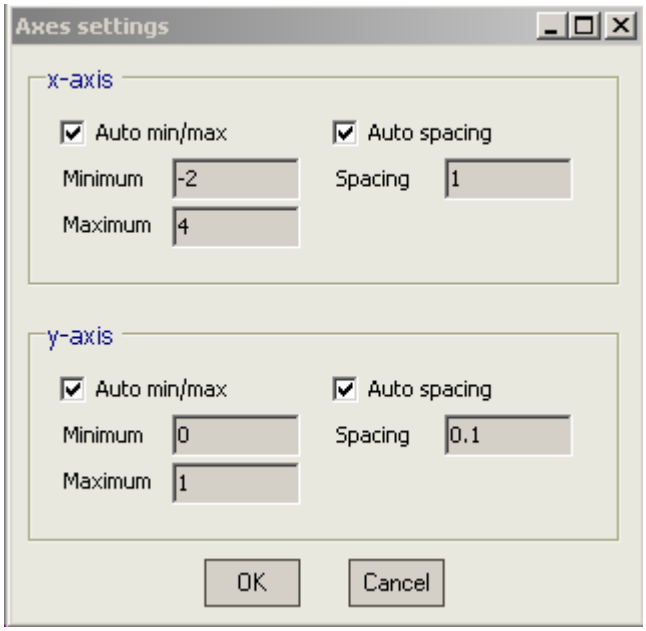


Figure 23. Axes settings dialog box.

Change the settings to your wishes and confirm by clicking **OK**.

5.9.3 Changing the default graph appearance

You may enter default values for the appearance of each of the three graphs via the **Tools**, **Preferences** menu that you can access via the menu bar. Figure 24 shows the dialog box that appears when you select this option.

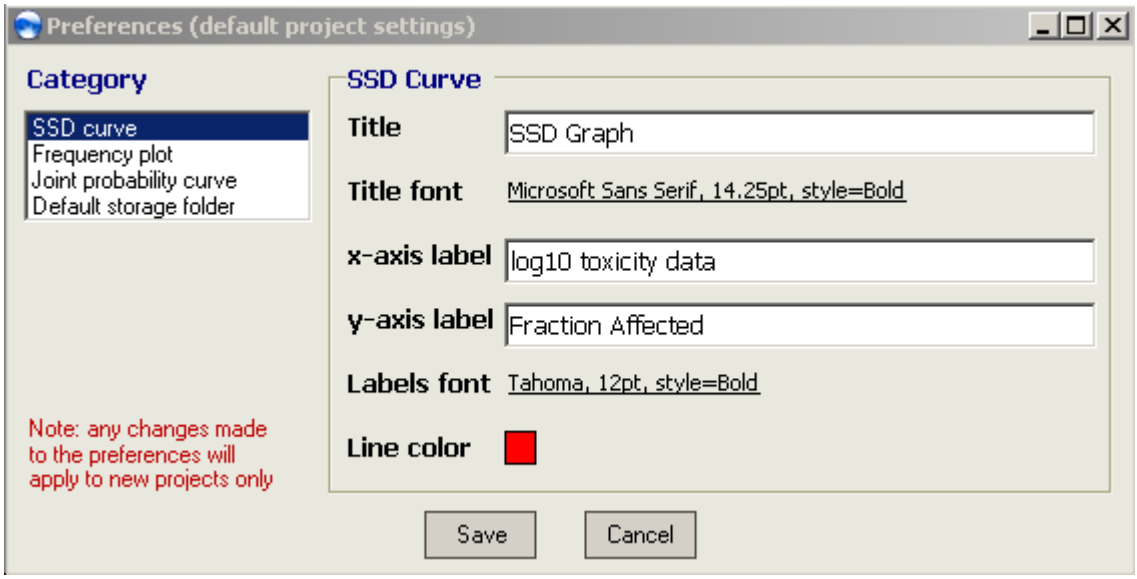


Figure 24. The default preferences dialog box.

- Select the graph of which you want to change the default settings in the box under **Category**.
- Edit the defaults in right half of the dialog box.
- After having made your changes, click on Save to save changes or on Cancel if you do not wish to keep your changes.
- Note that any changes you make will only apply to new projects that you will start after the current project. If you want to change the appearance of graphs in the current project, use the method described in the section above (5.9.1).

5.9.4 Changing the default file storage directory

You can select or change one directory where both your E_TX-project files as well as exported data files will be saved. To change this directory:

- Click on **Tools**, **Preferences**. The box shown in Figure 24 appears.
- Click on **Default storage folder** in the **Category** box. The box shown in Figure 25 appears.

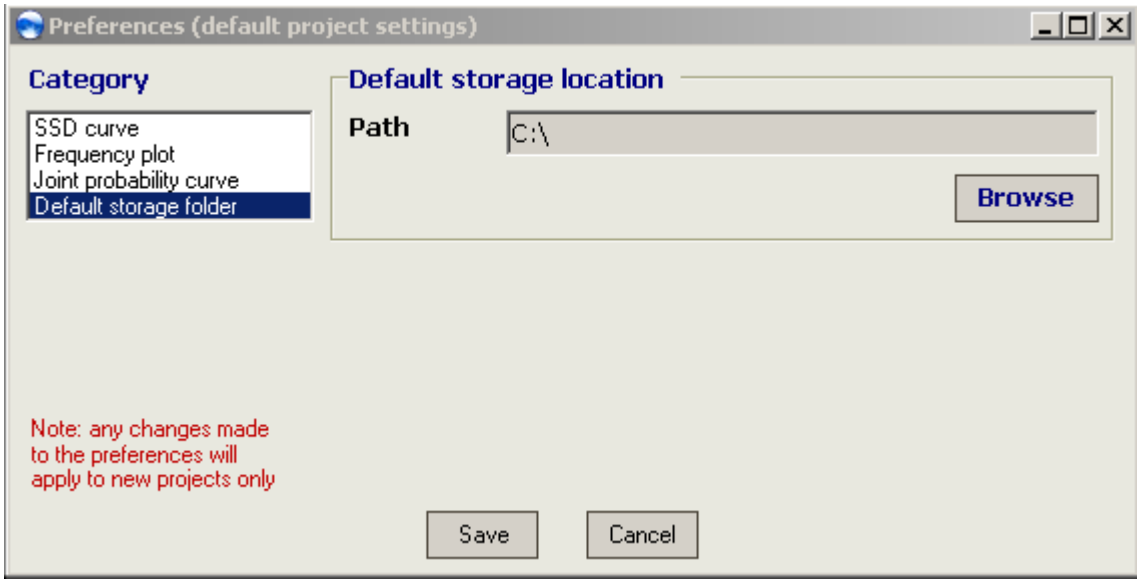


Figure 25. The default file storage location - dialog box.

- Type the directory of your choice behind **Path** or select a directory by browsing for it.
- Confirm your selection by clicking on **Save**. You will now leave the **Preferences** dialog.

5.10 Copying graphs

The three graphs that may be generated with *ETX* can be copied to other programs. This section explains how.

- Go to one of the graphs that you have generated: **SSD Histogram and PDF**, **Species sensitivity distribution** or **Joint probability curve** under **Graphics**.
- In the graph of your choice, click with the right mouse button (left button for left handed mousers).
- A pop up-window appears; select **Copy**.
- You can now paste your graph in another program.

5.11 Use of the decimal symbol

The default setting of *ETX* is that thousand separators will be ignored. In *ETX*, the regional settings of your computer are applied to the figures you enter in input cells and to the graphs that are displayed under **Graphics**. For help on changing of these regional settings, please see section 3.8.

The default setting of *ETX* has the following consequences. Example: if the period is your decimal symbol, the comma is usually the thousand separator. As long as you enter data with only a period as a decimal symbol, all is ok. You should enter figures higher than 999 *without* thousand separator: e.g. type 10000 rather than 10,000. As soon as *ETX* recognises a thousand separator, a warning will be displayed in the status bar at the bottom of your *ETX* screen: **Your data contains illegal decimal symbols**. This warning is displayed because otherwise, entering a figure like 13,1 would be entered as 131 and there is a fair chance that you might not notice this type of unexpected changes!

Therefore, two golden rules are:

1. **Do *not* use thousand separators when entering values.**
2. **The user determines the decimal symbol (settings) and is responsible for correct use.**

6. Limitations

6.1 SSD calculations

- For HC_5 and HC_{50} calculation a maximum number of 200 data can be entered.
- For calculation of the FA at the HC_{50} of your SSD (which is 50, by definition) and its accompanying lower and upper limit, the maximum number of toxicity data (*n*) is 75.

6.2 Calculation of the fraction affected

The calculation of the fraction affected that is calculated at a given environmental concentration (PEC) is limited by the height of the PEC value. The standardised logarithmic value of this PEC should be within the range -5 to 5. As a check, the value of the standardised logarithmic concentration calculated from your PEC value is shown in the **Fraction affected** output sheet. If this value is outside the range mentioned above, no results will be given. The message 'The method does not yield results because your standardised PEC exceeds the limits of <-5 or >5' will appear in red. The sheet will show 'Out of bounds!' as error values.

7. Examples

This section is a practical guide, which helps you through the calculation of an SSD or FA or other options in a stepwise manner.

7.1 Calculating an SSD

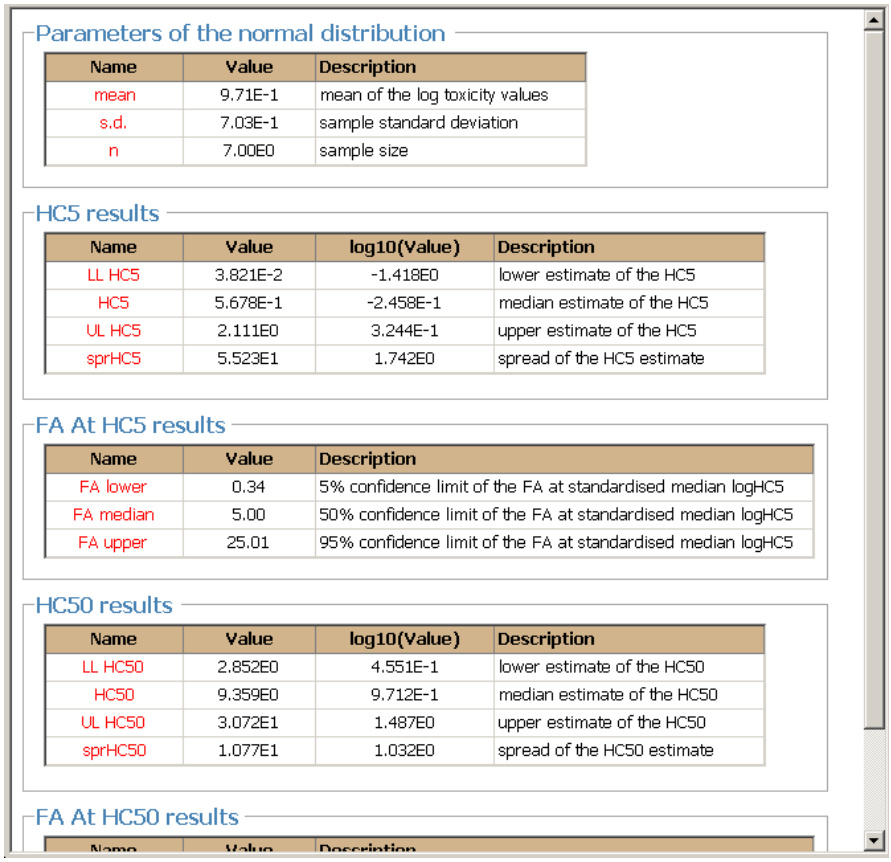
►Open *ETX* 2.0 or open a new project.
►In **Input data**, go to the **Input toxicity data** sheet, that is in the right part of your screen.
►Type the following data in the input cells in the column **Toxicity data**:

0.97
3.33
3.63
13.5
13.8
18.7
154

►In this example we will not add labels to the data. Leave the cells in the column **Label** empty.

These data are NOEC values for toxicity of cadmium to seven soil organisms, and are expressed in (µg Cd/g soil). The data are taken from Van Straalen and Denneman (1989) and are also used as an example in Aldenberg and Jaworska (2000).

►Go to the **Specifics** box in the **Input toxicity data** sheet. Click in the box to the right of **Unit**, and type: µg/g.
►In the **Specifics** box, click the pull down menu (downpointing arrow) and select NOEC.
►After having selected 'NOEC' in the **Specifics** box, press **Calculate, Go** in the Standard Menu bar. After performing the calculation, *ETX* automatically switches to the **Hazardous concentration sheet i**n the **Statistics section** (Figure 26, shown at the next page).
►Go to the **Goodness-of-fit** sheet and check if your data are normally distributed (Figure 27, shown at the next page). In this example it is probable that the data derive from a normal distribution.



Parameters of the normal distribution

Name	Value	Description
mean	9.71E-1	mean of the log toxicity values
s.d.	7.03E-1	sample standard deviation
n	7.00E0	sample size

HC5 results

Name	Value	log10(Value)	Description
LL HC5	3.821E-2	-1.418E0	lower estimate of the HC5
HC5	5.678E-1	-2.458E-1	median estimate of the HC5
UL HC5	2.111E0	3.244E-1	upper estimate of the HC5
sprHC5	5.523E1	1.742E0	spread of the HC5 estimate

FA At HC5 results

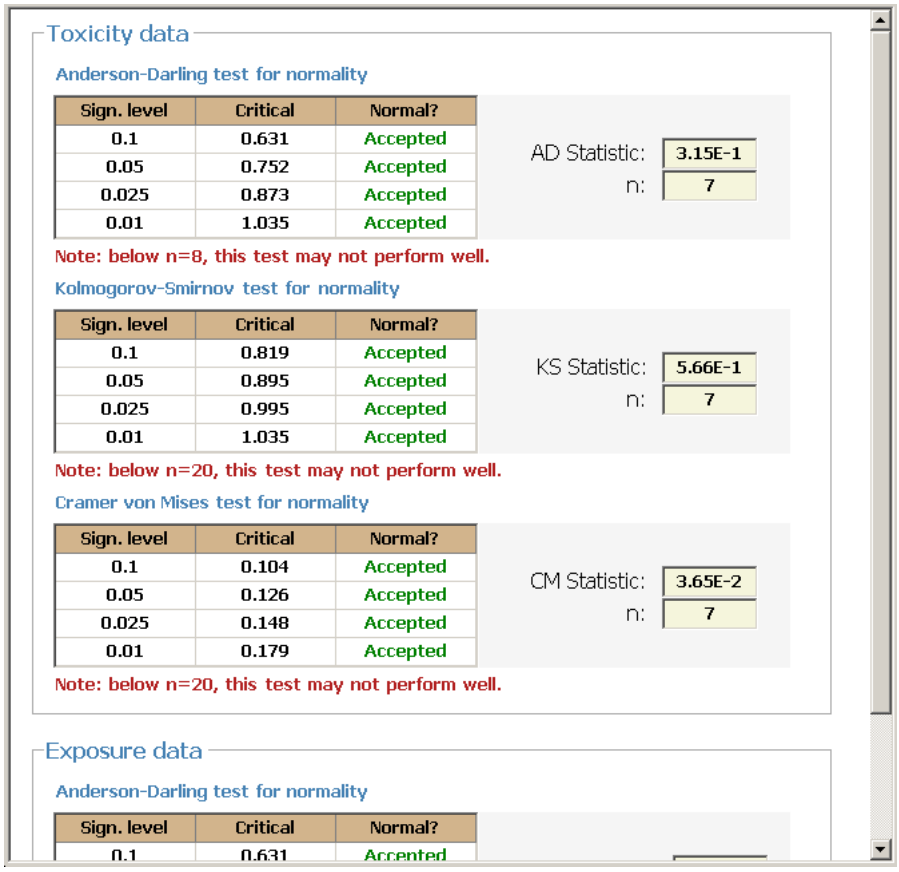
Name	Value	Description
FA lower	0.34	5% confidence limit of the FA at standardised median logHC5
FA median	5.00	50% confidence limit of the FA at standardised median logHC5
FA upper	25.01	95% confidence limit of the FA at standardised median logHC5

HC50 results

Name	Value	log10(Value)	Description
LL HC50	2.852E0	4.551E-1	lower estimate of the HC50
HC50	9.359E0	9.712E-1	median estimate of the HC50
UL HC50	3.072E1	1.487E0	upper estimate of the HC50
sprHC50	1.077E1	1.032E0	spread of the HC50 estimate

FA At HC50 results

Figure 26. Output of statistical parameters forHC5, HC_{50} and FA calculation.



Toxicity data

Anderson-Darling test for normality

Sign. level	Critical	Normal?
0.1	0.631	Accepted
0.05	0.752	Accepted
0.025	0.873	Accepted
0.01	1.035	Accepted

AD Statistic: 3.15E-1
n: 7

Note: below n=8, this test may not perform well.

Kolmogorov-Smirnov test for normality

Sign. level	Critical	Normal?
0.1	0.819	Accepted
0.05	0.895	Accepted
0.025	0.995	Accepted
0.01	1.035	Accepted

KS Statistic: 5.66E-1
n: 7

Note: below n=20, this test may not perform well.

Cramer von Mises test for normality

Sign. level	Critical	Normal?
0.1	0.104	Accepted
0.05	0.126	Accepted
0.025	0.148	Accepted
0.01	0.179	Accepted

CM Statistic: 3.65E-2
n: 7

Note: below n=20, this test may not perform well.

Exposure data

Anderson-Darling test for normality

Sign. level	Critical	Normal?
0.1	0.631	Accepted

Figure 27. Output goodness-of-fit tests.

►Go to **Output**, **Graphics** to find the **SSD Histogram and PDF** sheet and the **SSD graph** sheet to view the graphical output (Figure 28).

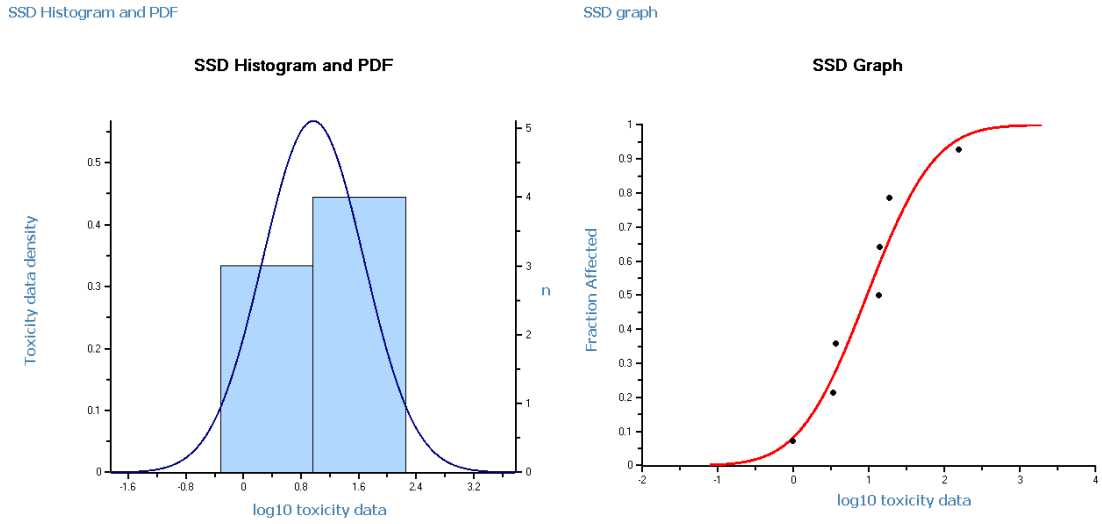
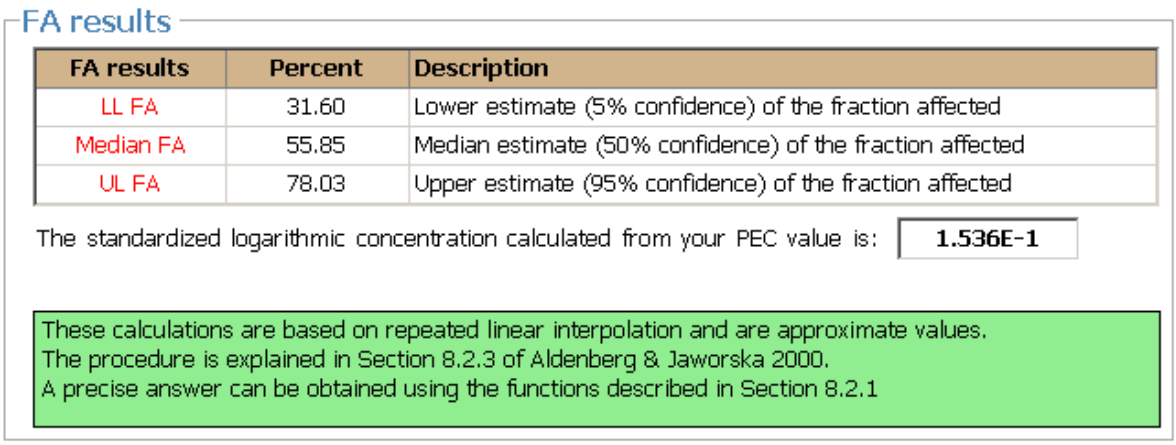


Figure 28. Histogram and PDF (left panel) and species sensitivity distribution (CDF, right panel).

7.2 Calculating a fraction of affected species

In order to calculate the fraction of species that is (potentially) affected at a given exposure concentration, we need the SSD of these species. So, you have to calculate an SSD first. In this example, we will use the SSD that was generated in section 7.1. Next, suppose we have one environmental concentration of 12 mg Cd/kg soil. The toxicity data were entered in the units of µg Cd/g soil. This is identical to mg Cd/kg soil. We have verified that units of toxicity data and PEC value are identical. Proceed as follows.

►Calculate an SSD (see section 7.1).
►Go to the **Input exposure data** sheet. In the right part of your screen, you see the box **Single PEC**.
►In the **Single PEC** box, click in the empty cell next to 'Enter a single PEC'; type 12 and press [enter].
►After having entered the PEC value, press **Calculate, Go** in the Standard Menu bar.
► ET_X will now switch automatically to the **Fraction affected** sheet in the **Output** section.
►Note that the value of the standardised logarithmic concentration calculated from this PEC is 0.5136. Since this value is >-5 and <5, an FA can be calculated by ET_X (see section 6 for limitations).
►The upper half of this sheet (the **FA results** box) shows your result (Figure 29): the median percentage of species in the soil that will be affected by this Cd concentration is 55.9%. The variation in the data tells you that there is 90% confidence that the percentage of affected species is between 31.6% and 78%.


FA results

FA results	Percent	Description
LL FA	31.60	Lower estimate (5% confidence) of the fraction affected
Median FA	55.85	Median estimate (50% confidence) of the fraction affected
UL FA	78.03	Upper estimate (95% confidence) of the fraction affected

The standardized logarithmic concentration calculated from your PEC value is: 1.536E-1

These calculations are based on repeated linear interpolation and are approximate values.
The procedure is explained in Section 8.2.3 of Aldenberg & Jaworska 2000.
A precise answer can be obtained using the functions described in Section 8.2.1

Figure 29. The fraction of species affected by 12 mg/kg Cd.

7.3 Calculating an expected ecological risk

Suppose you have a series of Cd concentrations, measured in a field soil. The measured concentrations are 1.5, 7, 3, 12, 1, 11 and 6 (mg Cd per kg soil). You can now calculate an EER and generate its graphical representation as a joint probability curve. In order to calculate the EER for the organisms in the field soil, we need the SSD of these species for Cd. In this example, we will use the SSD that was generated in section 7.1.

►Calculate an SSD (see section 7.1).
►Go to the **Input exposure data** sheet in the **Input data** section.
►In the right part of your screen, called **Input environmental concentrations**, click in the first cell of the column under **Exposure data**. Enter the following values, each next value in a new input cell:

1.5
7
3
12
1
11
6

►After entering the last value, press **Calculate, Go** in the Standard Menu bar.
►In order to check if your data follow a normal distribution, go to the **Goodness-of-fit** sheet in the **Output** section. In the lower half of this sheet, in the **Exposure data** box, the results of the test for normality of your exposure data are shown (Figure 30). In this example it is probable that the data derive from a normal distribution.

Expected Ecological Risk

FA results	Value	Description
SEC Mean	-4.882E-1	Scaled mean of (log10) environmental concentration distribution
SEC sd	5.996E-1	Scaled S.D. of (log10) environmental concentration distribution
Exp. Ecol. Risk	3.377E1	Expected Mean Ecological Risk, given the SSD and the EC distribution

The mean and standard deviation are scaled relative to the mean and SD of the Species Sensitivity Distribution.
The expected (mean) ecological risk is the area below the Joint Probability Curve.

Figure 30. Results of GOF test for exposure data.

►Go to the **Fraction affected** sheet (**Output**, **Statistics** section). In the lower half of this sheet (the **Expected ecological risk** box) you see two statistical parameters that are used to calculate the EER. The EER itself amounts to 33.8%, which is the percentage of species likely to be affected, given the SSD and EC distribution based on the data you have entered.

Expected Ecological Risk

FA results	Value	Description
SEC Mean	-4.882E-1	Scaled mean of (log10) environmental concentration distribution
SEC sd	5.996E-1	Scaled S.D. of (log10) environmental concentration distribution
Exp. Ecol. Risk	33.77	Expected Mean Ecological Risk, given the SSD and the EC distribution

The mean and standard deviation are scaled relative to the mean and SD of the Species Sensitivity Distribution.
The expected (mean) ecological risk is the area below the Joint Probability Curve.

Figure 31. Results of the expected ecological risk calculation.

►Go to the **Joint probability curve** sheet (**Output**, **Graphics** section). The graph in Figure 32 should be displayed.

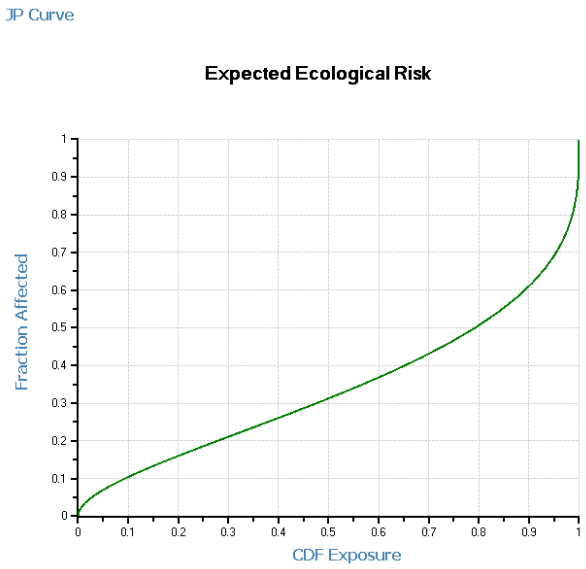


Figure 32. The joint probability curve belonging to the SSD from example 7.1 and the exposure data from example 7.3.

7.4 Calculating an HD_5 for birds or mammals from a small data set

In contrast to the previous examples, you do not have to generate an SSD before you can calculate an HD_5 for a small sample of toxicity data.

►Go to the **Input exposure data** sheet in the **Input data** section.
►In the **Input toxicity data** box (this is the right part of your screen), click on the **Use small sample approach** check box in the **Small sample** box. The following dialog box will appear:

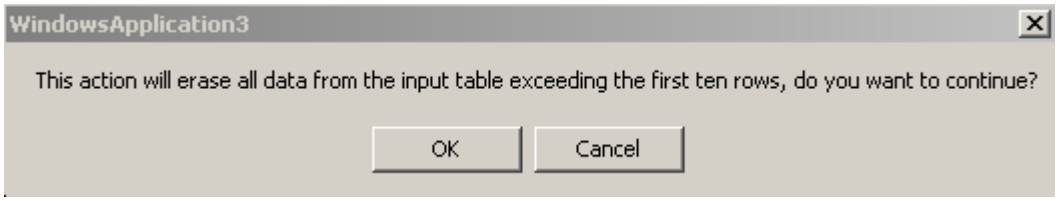


Figure 33. Small sample verification box.

If you had already typed data in the input section, these data will be lost when you switch to the Small sample input mode. Remember that this is a different method that does not require calculation of an SSD. You will have to start a separate project for each type of calculation that you perform. So, if you want to keep the data you had entered before, click on **Cancel** and save your work in an *ETX* project. Then start a new project for your 'small sample' calculation. Since we want to use the 'small sample' method, click on **OK**.
►In the part of your screen called **Input toxicity data**, the number of input cells is now reduced to 10.
►In the **Input toxicity data** sheet, click in the first cell of the column under **Toxicity data**. If there are any values present, delete these data. Type the following –fictitious– data in the input cells, each next value in a new input cell:

120
550
630

Let us say that these values are three LD_{50} values for birds, for a pesticide that is not a carbamate nor an organophosphorous compound.

►From the list of **Pre-defined standard deviations** that is also shown in the **Small sample** box, in the lower right corner of your screen, select the value for 'LD_{50} data of 55 pesticides for birds'. The figure 0.465 will appear in the **Standard deviation** cell.
►Click on **Calculate**, **Go** in the **Standard** menu bar.
►After performing the calculation, *ETX* automatically switches to the **Small sample output** sheet in the **Output**, **Statistics** section (Figure 34).

Parameters

Name	Value	Description
mean	2.540E0	mean of the log toxicity values
n	3	sample size
s.d.	0.465	preselected standard deviation

HD5 (median) results

Name	Value (mg/kg bwt)	Extrapolation factor (EF)
LL HD5	2.154E1	1.609E1
HD5 (median)	5.954E1	5.819E0
UL HD5	1.646E2	2.105E0

Figure 34. Output of the small sample *calculation with ETX 2.0.*

►The estimated 5th percentile of HD_5 values is 60 mg/kg_{bwt}. There is 90% confidence that the HD_5 will lie between 22 and 165 mg/kg_{bwt}. The extrapolation factors shown in the second column can be used to calculate each of the three reported parameters directly from the sample mean. For further details we refer to Aldenberg and Luttik (2002).

Acknowledgements

Thanks are due to ing. M. Adams, who was contact person at the Ministry of Housing, Spatial Planning and the Environment (VROM-DGM/SAS). We also thank Dr. E.M.J. Verbruggen, project manager of RIVM project 601501, for his contribution, Patrick Zweers for thorough auditing of the manual and Augustijn Onderzoek for their cooperative spirit. We gratefully acknowledge dr. M.P.M. Janssen, for coordinating the programming project and many helpful advices along the way.

We are also grateful to the people who kindly took part in the beta-testing procedure of E_TX: Christian Oberwalder (BASF AG), Michael Riffel (RIFCON), Magnus Wang (RIFCON), Elisabeth Erlacher (Universität für Bodenkultur), Margriet Beek (RIZA), Janny Pijnenburg (RIKZ), Ad Ragas (Radboud University), Cees van Gestel (Free University), Robert Luttik, Leo Posthuma, Eric Verbruggen, John Janus, Els Smit and Jean-Paul Rila (all RIVM). We greatly appreciate your efforts and many helpful comments!

The results as presented in this report have been discussed by the members of the scientific advisory group (WK-INS), who are acknowledged for their contribution. The advisory group provides a non-binding scientific comment on the final draft of a report in order to advise the steering committee of the project Setting (Inter) National Environmental Quality Standards (INS) on the scientific merits of the report.

References

Aldenberg T. 1993. *E_TX* 1.3a. A program to calculate confidence limits for hazardous concentrations based on small samples of toxicity data. Bilthoven, the Netherlands: National Institute of Public Health and Environmental Protection. Report no. 719102015. 52 pp.

Aldenberg T, Jaworska JS. 2000. Uncertainty of the hazardous concentration and fraction affected for normal species sensitivity distributions. Ecotoxicol Environ Saf 46: 1-18.

Aldenberg T, Jaworska JS, Traas TP. 2002. Normal Species Sensitivity Distributions and Probabilistic Ecological Risk Assessment. In: Posthuma L, Suter II GW, Traas TP, eds. Species Sensitivity Distributions in Ecotoxicology. Boca Raton, USA: Lewis Publishers. p. 49-102.

Aldenberg T, Luttik R. 2002. Extrapolation factors for tiny toxicity data sets from species sensitivity distributions with known standard deviation. In: Posthuma L, Suter II GW, Traas TP, eds. Species Sensitivity Distributions in Ecotoxicology. Boca Raton, USA.: Lewis Publishers. p. 103-118.

D'Agostino RB, Stephens MA. 1986. Goodness-of-Fit Techniques. New York, USA: Marcel Dekker, Inc. 560 pp.

Luttik R, Aldenberg T. 1997. Extrapolation factors for small samples of pesticide toxicity data: special focus on LD50 values for birds and mammals. Environ Toxicol Chem 16: 1785-1788.

Posthuma L, Suter II GW, Traas TP, eds. 2002. Species Sensitivity Distributions in Ecotoxicology. Boca Raton, USA: Lewis Publishers. 587 pp.

Scott DW. 1992. Multivariate density estimation. Theory, Practice and Visualization. New York, USA: John Wiley & Sons, Inc. 317 pp.

Van Straalen NM, Denneman CAJ. 1989. Ecotoxicological evaluation of soil quality criteria. Ecotoxicol Environ Saf 18: 241-251.

List of abbreviations

AUC	area under the curve
CDF	cumulative distribution function
DGM/SAS	Directorate General for Environmental Protection, Directorate for Chemicals, Waste and Radiation
EC	exposure concentration, in the context used here it is synonym to environmental concentration
ECD	exposure concentration distribution
EER	expected ecological risk
EU	European Union
E_TX	*E*co*To*X
FA	fraction affected
GOF	goodness-of-fit
HC_5	hazardous concentration, 5th percentile of normally distributed toxicity data
HC_{50}	hazardous concentration, median or 50th percentile of normally distributed toxicity data
HD_5	hazardous dose, 5th percentile of normally distributed toxicity data
INS	Setting (Inter)national Environmental Quality Standards
JPC	joint probability curve
LC_{50}	toxicant concentration causing 50% mortality in test population
LL	lower limit (of a confidence interval)
MS	Microsoft ©
n	sample size
NOEC	no observed effect concentration
PAF	potentially affected fraction; identical to FA, but PAF is no longer used in E_TX 2.0
PDF	probability density function
PEC	predicted environmental concentration
RIVM	National Institute for Public Health and the Environment
RIKZ	National Institute for Coastal and Marine management
RIZA	Institute Inland Water Management and Waste Water Treatment
s.d.	standard deviation
SEC	scaled environmental concentration
SSD	species sensitivity distribution
UL	upper limit (of a confidence interval)
VROM	Dutch Ministry of Housing, Spatial Planning and the Environment

Appendix 1 Installation problems

This appendix gives you an overview of combinations of Microsoft operating systems and Microsoft Office combinations we have tested plus the results of these tests. For some problems you may have benefit of a solution that is offered here.

Table 1. List of results: installation of E_TX 2.0 on MS operating systems and MS Office versions plus reference to possible solutions where appropriate.

MS Windows version	MS Office version	Test result	Solution
Windows 98 SE	not installed	Installation successful	
Windows 98	Office 97	Installation unsuccessful	
Windows 2000	Office 97	Installation unsuccessful	try solution 1
Windows 2000	2000	Installation successful	
Windows NT	not installed	Installation successful	
Windows NT	Office 97	Installation successful	
Windows XP	not installed	Installation successful	solution 2
Windows XP	Office 97	Installation successful	solution 2
Windows XP	Office 2000	Installation successful	solution 2
Windows XP	Office XP	Installation successful	solution 2
Windows XP	Office 2003	Installation successful	solution 2

Solution 1
With older versions of Windows, you might need to install Microsoft's Data Access Components (MDAC) or a recent update of MDAC. One version of MDAC is placed as a compressed file on the E_TX-CD. It is called **MDAC_TYP.exe** (version 2.80). You may install the file provided to you on the E_TX-CD, however, since updates of MDAC frequently appear we advise you to visit the Microsoft website to check for possible more recent updates. Visit: http://www.microsoft.com and search for MDAC.

If there are no MDAC installed on your PC or there are no recent updates available, follow the installation procedure described below.

Installation of MDAC from the E_TX-CD.
- Insert the E_TX-CD in the CD drive of your computer.
- Start the Windows Explorer (Start button, Programs, Accessories, Windows Explorer).
- In the left side of your screen, click on the drive that is your CD drive.
- In the right half of your screen, double click on the file **MDAC_TYP.exe**.
- The installation procedure will now start. Complete the installation procedure.

Now try to install E_TX. We refer to section 3.2 for the installation procedure.

Solution 2
If you have a machine with a recent Windows version, but installation is still unsuccessful, try manual installation of E_TX on your hard disk. Do as follows:

Manually Installing *ETX*

- Insert the *ETX*-CD in the CD drive of your computer.
- Start the Windows Explorer (Start button, Programs, Accessories, Windows Explorer).
- Create a new directory on your hard disk (e.g. C:\My Documents\ETX TEMP) in which the installation files can be placed.
- Files\RIVM*ETX* 2.0).
- Double click the file *ETX* 2.0.msi that is on the hard disk of your computer to start MS Installer.
- Normally, *ETX* should now be installed without problems.

Manually installing .NET framework

If the compact .NET framework is not installed on your computer, try to run the setup program manually:

- Insert the *ETX*-CD in the CD drive of your computer.
- In the dialog box enter E:\dotnetfx.exe, where E is the letter of your CD drive.
- Normally, the .NET framework should now be installed without problems.
- Next carry out the procedure described above, under 'Manually installing *ETX*'.

Appendix 2 Error messages

This appendix gives an overview of the error messages that can be displayed when using *E_TX* 2.0. A short description or solution to the problem is given where appropriate.

Error messages displayed during starting

Message

Could not initialize application, make sure a recent version of Microsoft's Data Access Components (MDAC) are installed on your PC

Solution

*S*ee solution 1 in Appendix 1.

Error messages displayed during calculations or entering data

Message

You should enter a set of toxicity data as well, in order to perform any calculations

Solution

E_TX can not perform calculation of an SSD and accompanying statistics (FA, HC etc.) unless you enter a toxicity data set in the **Input toxicity data** sheet. The minimum number of data is three.

The 'small sample' method can be used without calculating an SSD first.

Message

You haven't entered sufficient data!

Solution

E_TX will not perform calculation of an SSD and accompanying statistics (FA, HC etc.) unless you enter a toxicity data set of at least three values in the **Input toxicity data** sheet.

E_TX will not perform calculation of an HD_5 using the *small sample method* unless you enter at least one toxicity value in the **Input toxicity data** sheet.

Message

You haven't entered sufficient exposure data to calculate the Expected Ecological Risk and JPC curve

Solution

E_TX will not perform calculation of an EER and JPC unless you enter at least three exposure concentrations in the **Input exposure data** sheet.

Message

A label collection is active, this means you have to assign labels to all entries in the toxicity input screen before you can calculate results

Solution

Go to the **Input toxicity data** sheet and check if all toxicity entries have a label attached to it. If you do not want your toxicity data to be labelled, go to the **Input toxicity data** sheet and click on the **De-activate** button in the lower right corner of your screen.

Message
You have to enter a prefixed standard deviation in order to perform small sample calculations
Solution
In the **Small Sample** box in the **Input toxicity data** sheet, select a predefined standard deviation or enter a custom value.

Error messages displayed when saving an *ETX* file
Message
An error occurred while saving the file.
Solution
A single solution for this problem can not be given. Try copying the dataset or datasets you have entered to an other application. Then quit *ETX*, and restart *ETX* in order to retry your work.

Error messages displayed when opening an *ETX* file
Message
An error occurred while opening the file.
Solution
Your file may be damaged or in the wrong format. Try to open your data in another application and copy them to *ETX*.

Error messages displayed when exporting results
Message
Please (re)calculate your results before exporting
Solution
You have made changes to the data but you have not yet pressed the calculate button. Consequently, the results you are about to export do not belong to the dataset, which is currently entered in *ETX*. In order to continue, you have to update your results first by performing a calculation.

Message
An error occurred while exporting, please make sure that a version of Microsoft Excel 97 or higher is installed on your PC.
Solution
A solution is given in the message; an Excel version as recent as Excel 97 should be used.